QH
301
.B615
PR

Vol. 46, Parts 1 and 2 June 1959

UNIVERSITY OF MICHIGAN
JUL 2 - 1959
SCIENCE LIBRARY

BIOMETRIKA

FOUNDED BY
W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

MANAGING EDITOR
E. S. PEARSON

ASSOCIATE EDITOR
M. G. KENDALL

ISSUED BY
THE BIOMETRIKA OFFICE, UNIVERSITY COLLEGE, LONDON
PRINTED AT THE UNIVERSITY PRESS, CAMBRIDGE

[*Issued* 2 *June* 1959]

This volume of *Biometrika* is published with the co-operation of

F. N. DAVID	N. L. JOHNSON
J. DURBIN	D. G. KENDALL
J. B. S. HALDANE	D. V. LINDLEY
M. J. R. HEALY	R. L. PLACKETT

A volume containing about 500 pages will be published annually in two half-yearly issues, appearing in June and December.

Papers for publication should be sent to

PROFESSOR E. S. PEARSON

Department of Statistics, University College, London, W.C. 1

It is a condition of publication in *Biometrika* that the paper shall not already have been issued elsewhere, and will not be reprinted without leave of the Editors.

Contributors receive 25 copies of their papers free. Joint authors 15 copies each. Order forms for separates are sent to authors with proofs of their papers.

SUBSCRIPTIONS

The subscription price, payable in advance, is:

£2. 14*s.* (or **$8.00**) net per volume *including packing and postage*

BACK ISSUES

Volumes 20–45

These may be obtained from the BIOMETRIKA OFFICE, at the following prices, which include packing and postage:

Vols. 20–40: £5. 5*s.* (or $15.00) per volume; Vols. 41–45: £3. 10*s.* (or $10.00) per volume

Index (Subjects Vols. 1–37; Authors Vols. 1–43): 6*s.* (or $1.00)

Bound volumes: £1. 5*s.* (or $3.50) extra per volume; Binding cases: 12*s.* 6*d.* (or $1.75) each

Cheques should be made payable to *Biometrika*, crossed 'a/c BIOMETRIKA TRUST' and sent to THE SECRETARY, BIOMETRIKA OFFICE, UNIVERSITY COLLEGE, GOWER STREET, LONDON, W.C. 1, to whom all orders for series, single copies and offprints should be addressed.

Volumes 1–19

Permission for reprinting has been granted to Messrs Wm. Dawson & Sons Ltd. Volumes 1–19 are now ready for distribution, price £130 bound. Single vols. £7. 10*s.* each. Would librarians and others wishing to have copies please place their orders now with:

WM. DAWSON & SONS LTD.

16 WEST STREET, FARNHAM, SURREY, ENGLAND

VOLUME 46, PARTS 1 AND 2 JUNE 1959

ON THE CUMULANTS OF RENEWAL PROCESSES

BY WALTER L. SMITH†

DEDICATED TO THE MEMORY OF JOHN WISHART

1. INTRODUCTION

Let $X_1, X_2, X_3, \ldots$ be a *renewal process*; that is, a sequence of independent, non-negative, and identically distributed random variables which are not all zero with probability one. If there is a real, strictly positive $\tilde{\omega}$ such that, with probability one, all the X_i are divisible by $\tilde{\omega}$, then we shall call the renewal process *discrete*. A renewal process which is not discrete is termed *continuous*. Write $F(.)$ for the distribution function of X_1 and $\mu_r = \mathscr{E}X_1^r (\leqslant \infty)$, $S_k = X_1 + X_2 + \ldots + X_k$ and $F_k(.)$ for the distribution function of S_k.

Renewal theory is concerned primarily with the random variable N_t which is defined to be the maximum suffix k for which $S_k \leqslant t$, subject to the convention $N_t = 0$ if $X_1 > t$. In the obvious renewal application the X_i represent successive lifetimes of the object being renewed, and N_t is the number of renewals made by time t (subject to the original object having been installed at time 0). However, renewal theory has proved to be of significance in a wide range of investigations. We mention the theory of queues (Smith, 1953; Beneš, 1957), the theory of electronic counters (Takács, 1956; Smith, 1957), the theory of regenerative stochastic processes (Smith, 1955), the study of the variance-time curve, in its application to problems of nerve physiology (Cox & Smith, 1954; Cox, 1955), and the theory of cricket scores (Skellam & Shenton, 1957).

Previous theoretical work on renewal theory has been concerned almost exclusively (for very good reasons) with the first moment $\mathscr{E}N_t$. We refer to papers by Feller (1941), (1949), Täcklind (1945), Blackwell (1948), and Smith (1954). A typical result concerning $\mathscr{E}N_t$ is that if the renewal process is continuous and $\mu_2 < \infty$, then as $t \to +\infty$,

$$\mathscr{E}N_t = \frac{t}{\mu_1} + \left(\frac{\mu_2}{2\mu_1^2} - 1\right) + o(1).$$

This theorem was first proved, under slightly unnecessary restrictions, by Täcklind (1945), and in its final form by Smith (1954) (see also, Smith, 1957).

The study of higher moments of N_t seems to have been neglected, apart from a rather non-rigorous and short account by Skellam & Shenton (1957). Feller (1949) proves a theorem concerning the variance of N_t in the discrete renewal process, and Smith (1954) proves the corresponding result for continuous renewal processes: if $\mu_2 < \infty$ then as $t \to +\infty$

$$\operatorname{var} N_t = \frac{\mu_2 - \mu_1^2}{\mu_1^3} t + o(t).$$

A result of Cox & Smith (1954) concerning a related 'equilibrium' renewal process also suggests that if $\mu_3 < \infty$ then there are constants A and B such that as $t \to +\infty$, $\operatorname{var} N_t = At + B + o(1)$. A proof of such a result does not seem to be available, and nothing appears to have been established concerning the higher cumulants of N_t. It is with these higher cumulants that the present paper is concerned.

† This paper was prepared with the partial support of the Office of Naval Research.

The theorems concerning $\mathscr{E}N_t$ which are mentioned above have all been proved without irrelevant assumptions concerning $F(.)$. However, reference to Smith (1954) will show that a very special type of argument was needed to prove that $\operatorname{var} N_t \sim At$, which shows no prospect of generalizing to higher moments of N_t, nor of yielding the linear approximation $At+B$. Write $\mathscr{S}$ for the class of distribution functions $F(.)$ for which $F_k(.)$, for some finite k, has an absolutely continuous component. The present discussion of the higher cumulants of N_t, for continuous renewal processes, will be based on the weak, but extraneous, assumption that $F(.) \in \mathscr{S}$. *This assumption is assumed to hold throughout this paper even when not mentioned.* It will be noticed that once this extra assumption has been made the investigation proceeds with little real difficulty. Moreover, the fact that $F(.) \in \mathscr{S}$ will be found to imply certain variational properties of remainder terms that arise.

The paper falls naturally into two parts. The first part (§§ 2–5) is concerned with the rigorous proof that the nth cumulant of N_t has a linear asymptotic form, under general conditions. Two theorems (Theorems 3 and 4) are proved, which are jointly equivalent to the following.

THEOREM 0. *If $F(.) \in \mathscr{S}$ and $\mu_{n+p+1} < \infty$, $p \geqslant 0$, then there exist constants a_n and b_n such that the nth cumulant of N_t is given by*

$$a_n t + b_n + \frac{\lambda(t)}{(1+t)^p},$$

where $\lambda(t)$ is of bounded total variation, is $o(1)$ as $t \to +\infty$, satisfies the condition

$$\lambda(t) - \lambda(t-\alpha) = o(t^{-1})$$

as $t \to +\infty$ for every fixed $\alpha > 0$, and when $p \geqslant 1$ has the additional property that $\lambda(t)/(1+t)$ belongs to the class L_1.

This theorem substantially generalizes Theorems 8, 9, and 10 of Smith (1954) which were concerned with different aspects of Theorem 0 in the case $n = 1$.

The second part of the paper (§§ 6, 7) is concerned with the evaluation of a_n and b_n in terms of the μ_r. Their values are given for $n = 1, 2, \ldots, 8$.

In § 2 we introduce certain unconventional moments and cumulants which are more convenient in the present investigation than the conventional ones, and we study their properties. In § 3 a number of preliminary lemmas are proved, including in particular Lemma 4 which implies a property of the remainder term in Taylor expansions of characteristic functions which we believe to be of independent interest. In § 4 we develop a symbolic calculus for dealing with complicated Laplace transforms that arise and use this calculus to prove Theorems 1 and 2 concerning the asymptotic behaviour of moments of N_t. These results are preparatory to § 5 in which Theorem 0, above, is essentially proved.

In § 6 we show how the a_n and b_n, mentioned above, may be computed. The method we used is described in sufficient detail to make it possible to extend the list of formulae given for $\{a_n\}$ and $\{b_n\}$, if required. In § 7 various checking systems are discussed which assure us that the calculations of § 6 are correct. A special case which we will quote here, and which is a consequence of Theorem 0 and the computations in §§ 6 and 7, is as follows: If $\mu_3 < \infty$ then as $t \to +\infty$,

$$\operatorname{var} N_t = \frac{\mu_2 - \mu_1^2}{\mu_1^3} t + \left(\frac{5\mu_2^2}{4\mu_1^4} - \frac{2\mu_3}{3\mu_1^3} - \frac{\mu_2}{\mu_1^2}\right) + o(1).$$

The asymptotic linearity of the cumulants of N_t implies a curious combinatorial identity, which we have been unable to prove directly. We discuss this identity in an appendix. If it

could be established by a direct and elementary argument, then part of the proof of Theorem 0 could be drastically simplified.

Notice that this entire investigation is exclusively concerned with continuous renewal processes. Nevertheless, there is a parallel treatment for the discrete process, in which the requirement $F(.) \in \mathscr{S}$ can (and, of course, must) be dropped.

The author is grateful to his friend Dennis V. Lindley, who asked the questions which provoked the present work. He also wishes to thank Professor J. Neyman for making it possible for him to spend the summer of 1957 at Berkeley, where the bulk of this paper was prepared.

2. SOME MOMENT FORMULAE

Define the rth conventional moment $m_r(t) = \mathscr{E}N_t^r$ and write $k_r(t)$ for the corresponding conventional cumulant of order r. These conventional moments and cumulants are not the most suitable for a discussion of N_t. Instead, we shall use the moments

$$\phi_k(t) = \mathscr{E}(N_t+1)(N_t+2)\dots(N_t+k) \tag{2·1}$$

which come from the generating function

$$\Phi_t(\zeta) = \mathscr{E}\frac{1}{(1-\zeta)^{N_t+1}}.$$

The cumulants $\psi_k(t)$ are obtained by a formal power series expansion of the cumulant generating function $\Psi_t(\zeta) = \log \Phi_t(\zeta)$ in the usual way. Thus, the relations between the ϕ-moments and the ψ-cumulants are exactly similar to the corresponding relations between the conventional moments $m_r(t)$ and the conventional cumulants $k_r(t)$. It is evident, from (2·1), that

$$\begin{aligned}
\phi_1(t) &= m_1(t)+1,\\
\phi_2(t) &= m_2(t)+3m_1(t)+2,\\
\phi_3(t) &= m_3(t)+6m_2(t)+11m_1(t)+6,\\
\phi_4(t) &= m_4(t)+10m_3(t)+35m_2(t)+50m_1(t)+24.
\end{aligned}$$

In fact, it can easily be shown that

$$\phi_n(t) = |S_{n+1}^{n+1}|\, m_n(t) + |S_{n+1}^{n}|\, m_{n-1}(t) + |S_{n+1}^{n-1}|\, m_{n-2}(t) + \dots + |S_{n+1}^{1}|, \tag{2·2}$$

where S_n^m are Stirling's numbers of the first kind (Jordan, 1947, p. 142).

By systematic solution of equations (2·2) one can deduce the reciprocal relations

$$\begin{aligned}
m_1(t) &= \phi_1(t)-1,\\
m_2(t) &= \phi_2(t)-3\phi_1(t)+1,\\
m_3(t) &= \phi_3(t)-6\phi_2(t)+7\phi_1(t)-1,\\
m_4(t) &= \phi_4(t)-10\phi_3(t)+25\phi_2(t)-15\phi_1(t)+1.
\end{aligned}$$

However, it is not difficult to deduce from the defining property of the Stirling numbers, $\mathfrak{S}_n^m$, of the second kind (Jordan, 1947, p. 168) that the general relation is

$$m_n(t) = \mathfrak{S}_{n+1}^{n+1}\phi_n(t) - \mathfrak{S}_{n+1}^{n}\phi_{n-1}(t) + \dots + (-1)^n\mathfrak{S}_{n+1}^{1}. \tag{2·3}$$

Let us write $M_t(z) = \mathscr{E}e^{zN_t}$ for the conventional moment generating function, and $K_t(z) = \log M_t(z)$ for the corresponding cumulant generating function. It can be shown that $M_t(z) = e^{-z}\Phi_t(1-e^{-z})$ and hence that $K_t(z) = -z+\Psi_t(1-e^{-z})$. Thus we have

$$k_1(t)\,z + \frac{k_2(t)}{2!}z^2 + \dots = -z + \psi_1(t)[1-e^{-z}] + \frac{\psi_2(t)}{2!}[1-e^{-z}]^2 + \dots. \tag{2·4}$$

A result of Jordan (1947, p. 202) shows that

$$\frac{(1-e^{-z})^m}{m!} = \sum_{k=m}^{\infty} \frac{\mathfrak{S}_k^m (-1)^{m+k}}{k!} z^k; \tag{2·5}$$

and from (2·4) and (2·5) we deduce the following relations between conventional cumulants $k_n(t)$ and the ψ-cumulants $\psi_n(t)$.

$$\left.\begin{aligned} k_1(t) &= \psi_1(t) - 1, \\ k_n(t) &= \mathfrak{S}_n^n \psi_n(t) - \mathfrak{S}_n^{n-1} \psi_{n-1}(t) + \ldots + (-1)^{n-1} \mathfrak{S}_n^1 \psi_1(t) \quad (n \geqslant 2). \end{aligned}\right\} \tag{2·6}$$

For reference we list the following:

$$\left.\begin{aligned} k_2(t) &= \psi_2(t) - \psi_1(t), \\ k_3(t) &= \psi_3(t) - 3\psi_2(t) + \psi_1(t), \\ k_4(t) &= \psi_4(t) - 6\psi_3(t) + 7\psi_2(t) - \psi_1(t). \end{aligned}\right\} \tag{2·7}$$

Finally, to complete our discussion of the relations between the $k_n(t)$ and the $\psi_n(t)$, let us notice that

$$\Psi_l(\zeta) = \log(1-\zeta) + K_l(-\log[1-\zeta]). \tag{2·8}$$

If we use the following formula of Jordan ((1947), p. 202),

$$\frac{[\log(1+\zeta)]^m}{m!} = \sum_{n=m}^{\infty} \frac{S_n^m}{n!} \zeta^n,$$

then we can deduce from (2·8) that

$$\psi_r(t) = |S_r^r|\, k_r(t) + |S_r^{r-1}|\, k_{r-1}(t) + \ldots + |S_r^2|\, k_2(t) + |S_r^1|\, [k_1(t) - 1].$$

Whilst this last formula is occasionally useful for computational checks it will have no explicit use in the present paper.

3. Some notation and preliminary lemmas

All functions with which we shall be concerned will be assumed to vanish identically for negative values of their arguments. If $\Omega(t)$ is any function of bounded variation in every finite interval then we shall write

$$^*\Omega(s) = \int_{0-}^{\infty} e^{-st}\, d\Omega(t)$$

for the Laplace–Stieltjes transform (L.-S.T.) of $\Omega(t)$. If $\Omega(t)$ is any function which is integrable in every finite interval then we write

$$_*\Omega(s) = \int_0^{\infty} e^{-st} \Omega(t)\, dt$$

for the Laplace transform (L.T.) of $\Omega(t)$. The transforms that arise in the present work will all exist for $\mathscr{R}s > 0$. Notice that when $\Omega(t)$ has both a L.-S.T. and a L.T. then $^*\Omega(s) = s\,{}_*\Omega(s)$.

The following lemmas concerning $^*F(s)$ undoubtedly include much that is known. However, there does not seem to be available a satisfactory and complete treatment, dealing specifically with the one-sided Laplace transform, which leads to the somewhat surprising Lemma 4, below. For this reason we give the following proofs in some detail.

LEMMA 1. *If* $\mu_n < \infty$ *and*

$$\theta_n^s(x) = e^{-sx}\int_0^x e^{sz} z^n\, dz$$

then, as $s \to 0+$,

$$\int_{0-}^{\infty} \theta_n^s(x)\, dF(x) = o(s^{-1}).$$

Proof. For any $\epsilon > 0$ we can find an $x_0(\epsilon)$ such that

$$\int_{x_0}^{\infty} z^n\, dF(z) < \epsilon,$$

since $\mu_n < \infty$. But

$$\theta_n^s(x) < x^n \int_0^x e^{-sz}\, dz$$

$$< x^n/s.$$

Hence,

$$\int_{x_0}^{\infty} \theta_n^s(z)\, dF(z) < \epsilon/s. \tag{3·1}$$

Also, for $x \leqslant x_0$,

$$\theta_n^s(x) \leqslant \frac{x_0^n[1-e^{-sx_0}]}{s},$$

and for $s < s_0(x_0(\epsilon), \epsilon)$,

$$x_0^n[1-e^{-sx_0}] < \epsilon.$$

Thus

$$\int_{0-}^{x_0} \theta_n^s(z)\, dF(z) < \epsilon/s. \tag{3·2}$$

The lemma follows from (3·1) and (3·2).

LEMMA 2.

$$e^{-\alpha x} = 1 - \alpha x + \frac{\alpha^2 x^2}{2!} - \ldots + \frac{(-\alpha x)^n}{n!} + \frac{(-\alpha)^{n+1}}{n!}\theta_n^{\alpha}(x). \tag{3·3}$$

Proof. The L.T. of the right-hand side of (3·3) is

$$\frac{1}{s} - \frac{\alpha}{s^2} + \frac{\alpha^2}{s^3} - \ldots + \frac{(-\alpha)^n}{s^{n+1}} + \frac{(-\alpha)^{n+1}}{(s+\alpha)\, s^{n+1}} = \frac{1}{s+\alpha}.$$

LEMMA 3. *A necessary and sufficient condition for* $F(x)$ *to have its first* n *moments finite is that, for real* $s > 0$, $*F(s)$ *have an expansion*

$$*F(s) = 1 - \mu_1 s + \frac{\mu_2 s^2}{2!} - \ldots + \frac{(-s)^n \mu_n}{n!} + o(s^n).$$

Proof. Suppose first that $\mu_1, \mu_2, \ldots, \mu_n$ are finite. Then, by Lemmas 1 and 2 we have

$$*F(s) = \int_{0-}^{\infty} e^{-sz}\, dF(z)$$

$$= \int_{0-}^{\infty} \left\{1 - sz + \frac{s^2 z^2}{2!} - \ldots + \frac{(-s)^n}{n!} z^n + \frac{(-s)^{n+1}}{n!}\theta_n^s(z)\right\} dF(z)$$

$$= 1 - \mu_1 s + \frac{\mu_2}{2!} s^2 - \ldots + \frac{\mu_n}{n!}(-s)^n + s^{n+1} o\left(\frac{1}{s}\right).$$

This proves the necessity half of the theorem.

To prove the sufficiency part define the shift operator E^h by the rule $E^h g(s) = g(s+h)$, for any function $g(s)$. Then the kth derivative of $*F(s)$ *on the right* is given by

$$*F^{(k+)}(s) = \operatorname*{Lt}_{h=0+} \left(\frac{E^h - 1}{h}\right)^k *F(s).$$

If the expansion of $*F(s)$ announced in the statement of our lemma holds, then for $k \leqslant n$ we have

$$*F^{(k)}(0) = \mu_k < \infty.$$

However,
$$*F^{(k)}(s) = \operatorname*{Lt}_{h=0+} \left(\frac{E^h - 1}{h}\right)^k \int_{0-}^{\infty} e^{-sz}\, dF(z),$$

$$= \operatorname*{Lt}_{h=0+} \int_{0-}^{\infty} \left(\frac{e^{-hz} - 1}{h}\right)^k e^{-sz}\, dF(z).$$

Thus
$$\mu_k = \operatorname*{Lt}_{h=0+} \int_{0-}^{\infty} \left(\frac{e^{-hz} - 1}{hz}\right)^k z^k\, dF(z)$$

$$\geqslant \overline{\lim_{h=0+}} \frac{1}{2^k} \int_{0-}^{\epsilon h^{-1}} z^k\, dF(z),$$

where ϵ is chosen so that for all $0 \leqslant z \leqslant \epsilon$ we have $1 - e^{-z} \geqslant \frac{1}{2}z$. The finiteness of μ_k (regarded, temporarily, purely as a coefficient in the given expansion of $*F(s)$) establishes the finiteness of

$$\int_{0-}^{\infty} z^k\, dF(z).$$

LEMMA 4. *A necessary and sufficient condition for the distribution function $F(x)$ of a non-negative random variable to have its first k moments finite is that there exists another distribution function $F_{(k)}(x)$ such that*

$$*F(s) = 1 - \mu_1 s + \frac{\mu_2}{2!} s^2 - \ldots + \frac{\mu_{k-1}}{(k-1)!} (-s)^{k-1} + \frac{\mu_k}{k!} (-s)^k *F_{(k)}(s).$$

Proof. Obviously, if the above expansion is possible, then Lemma 3 and the fact that $*F_{(k)}(s) = 1 + o(1)$ as $s \to 0+$ prove the finiteness of the first k moments of $F(x)$.

Now suppose that we are merely given the fact that $\mu_k < \infty$. We shall assume $k \geqslant 1$ since the lemma is trivial if $k = 0$. If the mean μ_1 is finite then we can define a new distribution function

$$F_{(1)}(x) = \frac{1}{\mu_1} \int_0^x \{1 - F(z)\}\, dz,$$

with L.-S.T.
$$*F_{(1)}(s) = \frac{1 - *F(s)}{\mu_1 s}.$$

From Lemma 3 we see that

$$*F(s) = 1 - \mu_1 s + \frac{\mu_2}{2!} s^2 - \ldots + \frac{\mu_k}{k!} (-s)^k + o(s^k),$$

so that
$$*F_{(1)}(s) = 1 - \frac{\mu_2}{2!\,\mu_1} s + \frac{\mu_3}{3!\,\mu_1} s^2 - \ldots + \frac{\mu_k}{k!\,\mu_1} (-s)^{k-1} + o(s^{k-1}).$$

It then follows, also from Lemma 3, that the first $(k-1)$ moments of $F_{(1)}(x)$ are finite, and that the rth of these is given by $\mu_{r+1}/(r+1)\mu_1$. Notice, moreover, that

$$*F(s) = 1 - \mu_1 s *F_{(1)}(s).$$

We can now repeat the process, e.g. define

$$F_{(2)}(x) = \frac{2\mu_1}{\mu_2} \int_0^x \{1 - F_{(1)}(z)\}\, dz,$$

and obtain a new distribution function $F_{(2)}(x)$ whose first $(k-2)$ moments are finite, and given by the expansion

$$*F_{(2)}(s) = 1 - \frac{2\mu_3}{3!\,\mu_2}s + \frac{2\mu_4}{4!\,\mu_2}s^2 - \ldots + \frac{2\mu_k}{k!\,\mu_2}(-s)^{k-2} + o(s^{k-2}).$$

For this second 'derived' distribution function $F_{(2)}(x)$ we have

$$*F(s) = 1 - \mu_1 s + \frac{\mu_2}{2!}s^2 *F_{(2)}(s).$$

Evidently, repeated derivations of $F_{(3)}(x), F_{(4)}(x), \ldots, F_{(k)}(x)$, will prove the lemma.

Whilst strictly irrelevant to the present discussion, the remark is worth making that Lemma 4 is capable of trivial modification to include distributions of unrestricted (i.e. not necessarily positive) random variables. A final preparatory result is the following.

LEMMA 5. *If $\Omega(x)$ is any function of bounded (total) variation (B.V.) then*

$$*\Omega^{(k)}(s) = o(s^{-k}) \quad as \quad s \to 0+.$$

Proof.
$$(-s)^k *\Omega^{(k)}(s) = \int_{0-}^{\infty} e^{-sx}(sx)^k \, d\Omega(x).$$

But, as $s \to 0+$, $e^{-sx}(sx)^k \to 0$, boundedly. This proves the lemma.

4. THE ASYMPTOTIC FORM OF $\phi_n(t)$

We need first

LEMMA 6. $*\phi_n(s) = n!\{1 - *F(s)\}^{-n}$.

Proof. Since $\phi_n(t)$ is evidently a non-decreasing function of t it is legitimate to consider the L.-S.T. $*\phi_n(s)$.

Now
$$\phi_n(t) = \mathscr{E}(N_t+1)(N_t+2)\ldots(N_t+n)$$
$$= \sum_{k=0}^{\infty}(k+1)(k+2)\ldots(k+n)\{F_k(t) - F_{k+1}(t)\}.$$

Thus
$$*\phi_n(s) = \sum_{k=0}^{\infty}(k+1)(k+2)\ldots(k+n)\{[*F(s)]^k - [*F(s)]^{k+1}\}$$
$$= n!\{1 - *F(s)\}\sum_{k=0}^{\infty}\frac{(k+1)\ldots(k+n)}{n!}[*F(s)]^k.$$

The lemma follows at once.

We remark that it is the particularly simple form of the L.-S.T. $*\phi_n(s)$ that makes $\phi_n(t)$ so much more convenient, for our present study, than the conventional moment $m_n(t)$.

Provided $\mu_n < \infty$ we can consider the function

$$*K_n(s) = \frac{1 - *F_{(n)}(s)}{1 - *F(s)} \quad (n \geqslant 1),$$

concerning which we have the following three lemmas

LEMMA 7. *If $F(t) \in \mathscr{S}$ and $\mu_{n+1} < \infty$ then $K_n(t)$ is of bounded total variation, and*

$$K_n(\infty) = \frac{\mu_{n+1}}{(n+1)\mu_1\mu_n}.$$

Proof. The fact that $K_n(t)$ is of bounded total variation follows from the fact that if $\mu_{n+1} < \infty$ then the means of both $F(t)$ and $F_{(n)}(t)$ are finite, and from Lemma 6 of Smith (1954). The mean of $F_{(n)}(t)$ is found, from the Taylor expansion of $*F_{(n)}(s)$, to be $\mu_{n+1}/(n+1)\mu_n$, and so, letting $s \to 0+$ in $*K_n(s)$ we deduce that $*K_n(0) = K_n(\infty)$ has the value announced.

LEMMA 8. *If $\mu_{n+1} < \infty$, so that $^*K_{n+1}(s)$ exists, then*

$$\frac{^*K_n(s)}{s} = \frac{\mu_{n+1}}{(n+1)\mu_n}\left[\frac{1}{\mu_1 s} + {}^*K_1(s) - {}^*K_{n+1}(s)\right]. \tag{4·1}$$

Proof. We know that

$$^*F_{(n)}(s) = 1 - \frac{\mu_{n+1}}{(n+1)\mu_n} s\,{}^*F_{(n+1)}(s).$$

But the right-hand side of (4·1) is equal to

$$\frac{\mu_{n+1}}{(n+1)\mu_n}\left[\frac{^*F_{(1)}(s)}{1-{}^*F(s)} + \frac{1-{}^*F_{(1)}(s)}{1-{}^*F(s)} - \frac{1-{}^*F_{(n+1)}(s)}{1-{}^*F(s)}\right] = \frac{\mu_{n+1}}{(n+1)\mu_n}\frac{^*F_{(n+1)}(s)}{1-{}^*F(s)}.$$

The lemma is thus evident.

LEMMA 9.

$$^*\phi_1(s) = \frac{1}{\mu_1 s} + {}^*K_1(s).$$

Proof. Direct computation establishes the lemma.

So far we have been able to deal exactly and precisely with the functions and coefficients which have arisen. If we were to attack in this way the theorems which we now approach we would rapidly become involved in hopeless complication. Our primary objective in this section is merely to establish that $\phi_n(t)$ has a polynomial form with a remainder term of a certain type. At this stage we shall therefore introduce, and make considerable use of, certain notations which will preserve the essence of the various relations in which we have interest, without preserving a great deal of undesired and heavy detail. We shall write γ for any constant, and γ_n for any finite rational function of $\mu_1, \mu_2, \ldots, \mu_n$. For instance, we would replace $\mu_1^2\mu_3$ by γ_3. However, nothing is to stop us calling $\mu_1^2\mu_3$ by the name γ_{12}, or γ_{13}, if we want to. We shall also write $^*M_w^d$ for any multinomial expression in $^*K_1, {}^*K_2, \ldots, {}^*K_w$, with coefficients γ_w and with the maximum degree of any term not greater than d. We shall also insist, when $w > 1$, that the degree of $^*K_w$ in any term is not to exceed unity. For examples of the use of this symbolism we offer the following.

(i) $$\mu_1^4\mu_4\,{}^*K_3\,{}^*K_4 + \mu_2\,{}^*K_2^3 = {}^*M_4^3,$$

(ii) $$\mu_1\,{}^*K_6\,{}^*K_1^5 + \mu_6\,{}^*K_5 = {}^*M_6^6.$$

Certain points about the notation $^*M_w^d$ should be noticed. The multinomial terms in $^*M_w^d$ do not necessarily have to be of homogeneous degree. A $^*K$ or γ with maximum allowable suffix does not necessarily have to occur. The essential feature of the notation $^*M_w^d$ is that if $\mu_{w+1} < \infty$ then it represents the L.-S.T. of a function of B.V. This is because any γ_w is certainly defined and finite if $\mu_{w+1} < \infty$, every $^*K_r(s)$ with $r \leqslant w$ will be (by Lemma 7) the L.-S.T. of a function of B.V., and the product of a finite number of L.-S.T.'s of functions of B.V. is itself the L.-S.T. of a function of B.V. Incidentally, it is easy to see from Lemma 7 that $M_w^d(+\infty)$, which is the same thing as $^*M_w^d(0)$, the total net variation of $M_w^d(t)$, is simply γ_{w+1}.

Expressions such as γ_3, $^*M_4^2$, etc., may occur several times in the same equation; but it must not be supposed that at each appearance they represent the same function. However,

the following laws of combination should be noticed. If $r = \max(p, q)$ and $d = \max(b, c)$, then

$$\gamma_p + \gamma_q = \gamma_r, \tag{C1}$$

$$\gamma_p \gamma_q = \gamma_r, \tag{C2}$$

$$*M_p^b + *M_q^c = *M_r^d, \tag{C3}$$

$$\left.\begin{aligned} *M_p^b *M_q^c &= *M_{r+1}^{b+c} \quad \text{if} \quad p = q \neq 1 \\ &= *M_r^{b+c} \text{ otherwise.} \end{aligned}\right\}, \tag{C4}$$

The slight complication in rule (C4) arises from the fact that in a multinomial $*M_w^d$ with $w > 1$, a $*K_w$ may not have a higher degree than unity in any term. When we multiply two multinomials, of the same weight w, it is possible for a term containing a factor $*K_w^2$ to arise.

It should be clear that we may consistently write

$$*M_w^0 = \gamma_w.$$

LEMMA 10. *If* $d \geqslant 1$, *then*

$$\frac{*M_w^d}{s} = \frac{\gamma_{w+1}}{s} + *M_{w+1}^d.$$

Proof. A typical term in $*M_w^d$ could be written as $\gamma_w *K_{p_1} *K_{p_2} \ldots *K_{p_r}$, where $r \leqslant d$ and every $p_j \leqslant w$. This term could be written, alternatively, as $*K_{p_1} *M_w^{d-1}$. But Lemma 8 states, essentially, that

$$\frac{*K_{p_1}}{s} = \frac{\gamma_{p_1+1}}{s} + *M_{p_1+1}^1.$$

Thus, remembering that $p_1 \leqslant w$, we have by (C4) that

$$\frac{*K_{p_1} *M_w^{d-1}}{s} = \frac{\gamma_{w+1} *M_w^{d-1}}{s} + *M_{w+1}^d.$$

From this symbolic reduction of a typical term in $*M_w^d$ it is thus apparent that

$$\frac{*M_w^d}{s} = \gamma_{w+1} \frac{*M_w^{d-1}}{s} + *M_{w+1}^d. \tag{4·2}$$

If $d = 1$ then the lemma is proved, since $*M_w^0 = \gamma_w$. If not, we re-apply the above argument to $*M_w^{d-1}$, and deduce

$$\frac{*M_w^{d-1}}{s} = \gamma_{w+1} \frac{*M_w^{d-2}}{s} + *M_{w+1}^{d-1}. \tag{4·3}$$

On combining (4·2) and (4·3), and using (C2) and (C3), we have

$$\frac{*M_w^d}{s} = \gamma_{w+1} \frac{*M_w^{d-2}}{s} + *M_{w+1}^{d-1} + *M_{w+1}^d = \gamma_{w+1} \frac{*M_w^{d-2}}{s} + *M_{w+1}^d. \tag{4·4}$$

If $d = 2$ the lemma is proved at this stage, and it should now be clear that repeated reductions will eventually prove the lemma, whatever the value of d.

In our discussions of remainder terms we shall find it convenient to use the following classes of functions, which we now define. All functions, as has been stated before, are assumed to vanish for negative argument.

DEFINITION 1. The function $\lambda(t)$ belongs to the class $\mathscr{B}$ if and only if it is B.V., tends to zero as t approaches $+\infty$, and satisfies the condition

$$\lambda(t)-\lambda(t-\alpha)=o(t^{-1}), \tag{*}$$

as $t\to+\infty$, for every $\alpha>0$.

DEFINITION 2. The function $\lambda(t)$ belongs to the class $\mathscr{L}$ if and only if $\lambda(t)/(1+t)$ belongs to L_1.

It will be convenient to write $\mathscr{R}=\mathscr{B}\cap\mathscr{L}$ for the class of all those functions in both $\mathscr{B}$ and $\mathscr{L}$. An easy result which we use later is

$$\text{If}\quad \lambda_1(t)\in\mathscr{B}\quad\text{and}\quad\lambda_2(t)\in\mathscr{R},\quad\text{then}\quad\lambda_1(t)\,\lambda_2(t)\in\mathscr{R}. \tag{C 5}$$

(In fact $\mathscr{B}$ is a commutative algebra over the reals and (C 5) is included in the remark that $\mathscr{R}$ is an *ideal* in $\mathscr{B}$. However, these concepts will find no application in the present work.)

LEMMA 11. *If* $\mu_w<\infty,\quad w>1$, *then*

(i) $M_w^d(t)=o(t)$ *as* $t\to\infty$,

(ii) *whenever* $G(t)$ *is B.V., the function* $P(t)$ *whose L.-S.T. is* $*G*M_w^d$ *satisfies the conditions*

$$P(t)-P(t-\alpha)=o(1)\quad as\quad t\to\infty,\quad for\ every\quad\alpha>0. \tag{**}$$

Proof. Terms in $*M_w^d$ may contain $*K_w$ to at most the first power so that we can evidently write $*M_w^d$ as $*K_w*J+*L$, where $*J$ and $*L$ contain only $*K_p$ terms with $p<w$. Since $\mu_w<\infty$, it follows that $J(t)$ and $L(t)$ are both functions of B.V. Moreover, if we write $*\Psi=*J\{1-*F_{(w)}\}$ then $\Psi(t)$ is also a function of B.V. and is, indeed, one which vanishes at $+\infty$ (since $*\Psi(0+)$ is obviously zero). Rewriting $*M_w^d$ as $*L+*\Psi*\phi_1$ we see that

$$M_w^d(t)=L(t)+\int_0^t\Psi(t-t')\,d\phi_1(t'),$$

and the fact that $M_w^d(t)=o(t)$ now follows from Lemma 1 of Smith (1955).

To prove the second part of the lemma we notice that the L.-S.T. of $P(t)-P(t-\alpha)$ is simply

$$\{1-e^{-\alpha s}\}*G*M_w^d=\{1-e^{-\alpha s}\}*G*L+\frac{1-e^{-\alpha s}}{1-*F}*G*\Psi,$$

which we recognize, by Lemma 6 of Smith (1954), to be the L.-S.T. of a function of B.V. Evidently this function approaches zero as $t\to+\infty$, as may be seen by letting $s\to0+$. Thus the lemma is proved.

LEMMA 12. *For any non-negative integer* p, *define* $*g_p(s)=s^{-p}*M_w^d$. *Then, if* $\mu_{w+1}<\infty$, $w\geqslant1$, *we can find a* γ_{w+1} *such that* $t^{-p}g_p(t)-\gamma_{w+1}$ *belongs to* $\mathscr{B}$.

Proof. Evidently

$$g_p(t)=\int_0^t(t-z)^p\,dM_w^d(z),$$

where M_w^d is B.V. Hence

$$\frac{g_p(t)}{t^p}=M_w^d(t)+\sum_{r=1}^{r=p}\int_0^t\left(\frac{z}{t}\right)^r dM_w^d(z).$$

This suggests that we take for the γ_{w+1} referred to in the enunciation the value $M_w^d(+\infty)$. Since $M_w^d(t)$ is B.V. we shall have proved that $t^{-p}g_p(t)-M_w^d(+\infty)$ is B.V., and $o(1)$ at $+\infty$, if we can prove that

$$l_r(t)=\frac{1}{t^r}\int_0^t z^r dV(z)$$

is B.V. and vanishes at $+\infty$, for every integer $r>0$, and every distribution function $V(z)$.

Choose $\epsilon > 0$, and take T_1 so that $V(T_1) > 1-\epsilon$. Then for all $t > T_2 > \epsilon^{-1} T_1$ it is clear that

$$\begin{aligned} l_r(t) &< \frac{1}{T_2^r}\int_0^{T_1} z^r dV(z) + \int_{T_1}^t dV(z) \\ &< \epsilon^r + \epsilon. \end{aligned}$$

Thus $l_r(t)$ does in fact vanish at $+\infty$. Moreover, an integration by parts will quickly verify that

$$l_r(t) = V(t) - r\int_0^t \frac{l_r(z)}{z}\,dz.$$

But, because $l_r(t)$ is non-negative, this displays $l_r(t)$ to be the difference between two increasing functions which must, since $V(t) \leqslant 1$, be bounded. Hence $l_r(t)$ is B.V.

To finish the proof we must show that the condition (*) used in Definition 1 is satisfied by $t^{-p} g_p(t)$. To this end we remark that an easy deduction from Lemma 10 is that

$$\frac{{}^*M_w^d}{s^p} = \frac{M_w^d(+\infty)}{s^p} + \frac{{}^*M_{w+1}^d}{s^{p-1}}.$$

Hence

$$\frac{g_p(t)}{t^p} - M_w^d(+\infty) = \frac{1}{t}\int_0^t \left(1-\frac{z}{t}\right)^{p-1} dM_{w+1}^d(z), \tag{4·5}$$

where $M_{w+1}^d(z)$ cannot be assumed to be B.V. However, let us define

$$m(t) = \int_0^t \left(1-\frac{z}{t}\right)^{p-1} dM_{w+1}^d(z). \tag{4·6}$$

We claim that the proof will be complete if we prove that $m(t)$ satisfies the condition

$$m(t) - m(t-\alpha) = o(1) \quad \text{as} \quad t \to +\infty, \quad \text{for every } \alpha > 0. \tag{**}$$

For, by (4·5) and the first part of this proof, we have already shown that $m(t) = o(t)$ as $t \to +\infty$. Thus $\{1 - t/(t-\alpha)\}\, m(t-\alpha) = o(1)$, and so, if (**) holds then (*) is satisfied by $m(t)/t$ (which is the result required).

To reduce our problem yet further we can now see from (4·6) that all we need prove is that if, for any $r \geqslant 0$, we define

$$n_r(t) = \frac{1}{t^r}\int_0^t z^r dM_{w+1}^d(z),$$

then $n_r(t)$ satisfies (**).

An integration by parts (identical with one performed already in this proof) shows that

$$\frac{n_r(t)}{t} + \frac{r}{t}\int_0^t \frac{n_r(z)}{z}\,dz = \frac{M_{w+1}^d(t)}{t}. \tag{4·7}$$

Lemma 11 (i) has proved that the right-hand side of (4·7) tends to zero as $t \to +\infty$, and so we can deduce from the integral version of Mercer's theorem (Paley & Wiener (1934) p. 58) that $n_r(t) = o(t)$ as $t \to +\infty$. Hence

$$\lim_{t=+\infty} \int_{t-\alpha}^t \frac{n_r(z)}{z}\,dz = 0. \tag{4·8}$$

Multiply (4·7) by t. Then the fact that $n_r(t)$ satisfies (**) follows from the resulting equation, (4·8), and the result that $M_{w+1}^d(t)$ satisfies (**), which was proved in Lemma 11 (ii).

THEOREM 1. *If* $\mu_{n+1} < \infty$ *then*

$$\phi_n(t) = \gamma_1 t^n + \gamma_2 t^{n-1} + \ldots + \gamma_n t + \gamma_{n+1} + \omega(t),$$

where $\omega(t) \in \mathscr{B}$.

Proof. By Lemma 6

$$\begin{aligned} {}^*\phi_n(s) &= n! \, [{}^*\phi_1(s)]^n \\ &= n! \left[\frac{\gamma_1}{s} + {}^*K_1\right]^n, \quad \text{by Lemma 9,} \\ &= \frac{\gamma_1}{s^n} + \frac{{}^*M_1^1}{s^{n-1}} + \frac{{}^*M_1^2}{s^{n-2}} + \ldots + {}^*M_1^n. \end{aligned}$$

(It is to be recalled that the rule excluding powers of ${}^*K_w$ greater than unity in ${}^*M_w^d$ does not apply when $w = 1$.)

Let us now apply Lemma 10 repeatedly. We deduce that

$$\begin{aligned} {}^*\phi_n(s) &= \frac{\gamma_1}{s^n} + \frac{\gamma_2}{s^{n-1}} + \frac{{}^*M_2^2}{s^{n-2}} + \ldots + {}^*M_2^n, \\ &= \frac{\gamma_1}{s^n} + \frac{\gamma_2}{s^{n-1}} + \frac{\gamma_3}{s^{n-2}} + \frac{{}^*M_3^3}{s^{n-3}} + \ldots + {}^*M_3^n, \\ &= \text{etc.}, \\ &= \frac{\gamma_1}{s^n} + \frac{\gamma_2}{s^{n-1}} + \ldots + \frac{\gamma_n}{s} + {}^*M_n^n. \end{aligned} \tag{4·9}$$

Thus it follows that

$$\phi_n(t) = \gamma_1 t^n + \gamma_2 t^{n-1} + \ldots + \gamma_n t + M_n^n(t).$$

The property claimed for the remainder $\omega(t)$ follows immediately from Lemma 12, in which we let $p = 0$.

LEMMA 13. *If* $\mu_{n+p} < \infty$, $n > 0$, $p > 0$, *then*

$$(-1)^p \frac{{}^*F^{(p)}(s)}{\mu_p}$$

is the L.-S.T. of a distribution function $F^{[p]}(x)$, *say, whose first n moments are finite.*

Proof. Evidently

$$(-1)^p \, {}^*F^{(p)}(s) = \int_{0-}^{\infty} e^{-sx} x^p \, dF(x),$$

and so, if we put

$$dF^{[p]}(x) = \frac{x^p \, dF(x)}{\mu_p}$$

the lemma is obvious.

THEOREM 2. *If* $\mu_{n+p+1} < \infty, \quad n > 0, \quad p > 0$, *then*

$$\phi_n(t) = \gamma_1 t^n + \gamma_2 t^{n-1} + \ldots + \gamma_{n+1} + \frac{\lambda(t)}{(1+t)^p},$$

where $\lambda(t) \in \mathscr{R}$.

Remark. This theorem bears a close superficial resemblance to the previous theorem, but provides us with much more information about the asymptotic behaviour of the remainder term, and requires a somewhat more intricate proof. We need the extra information concerning the remainder term if we are to discuss the asymptotic behaviour of the cumulants, which involve *products* of moments.

Proof. The proof is based on the fact that $(-1)^p {}_*\phi_n^{(p)}(s)$ is the L.T. of $t^p\phi_n(t)$. However, before we enter the main part of this proof we shall establish the preliminary result

$$(-1)^p {}_*\phi_n^{(p)}(s) = \frac{\gamma_1}{s^{n+p+1}} + \frac{\gamma_2}{s^{n+p}} + \ldots + \frac{\gamma_{n+1}}{s^{p+1}} + o\left(\frac{1}{s}\right). \tag{4·10}$$

To prove this preliminary result, we observe that Lemma 10 leads easily to

$$\frac{{}^*M_n^n}{s} = \frac{\gamma_{n+1}}{s} + \gamma_{n+2} + \gamma_{n+3}s + \ldots + \gamma_{n+p}s^{p-2} + s^{p-1}{}^*M_{n+p}^n. \tag{4·11}$$

However, Lemma 5 and some repeated differentiation will demonstrate that

$$\left(-\frac{d}{ds}\right)^p s^{p-1}{}^*M_{n+p}^n = o\left(\frac{1}{s}\right),$$

so that we can deduce from (4·11) the fact that

$$\left(-\frac{d}{ds}\right)^p \frac{{}^*M_n^n}{s} = \frac{\gamma_{n+1}}{s^{p+1}} + o\left(\frac{1}{s}\right). \tag{4·12}$$

The preliminary result (4·10) then follows at once from (4·9) and (4·12).

Turning now to the main part of the proof we remark that by Lemma 6 we can write

$${}_*\phi_n(s) = \frac{n!}{s}\left\{\frac{1}{1-{}^*F(s)}\right\}^n,$$

so that $(-1)^p {}_*\phi_n^{(p)}(s)$ is the sum of a number of terms like

$$\Gamma = \gamma \frac{{}^*F^{(j_1)}(s)\ldots{}^*F^{(j_r)}(s)}{s^{t+1}\{1-{}^*F(s)\}^{n+r}}.$$

In such a term, Γ, the number of factors in the numerator is r, which is also the increase above n of the power of $\{1-{}^*F(s)\}$ in the denominator. The multiplier γ will depend only on n, and not on the moments $\mu_1, \mu_2, \ldots$ etc. Furthermore, one must have

$$j_1 + j_2 + \ldots + j_r + t = p,$$

so that
$$r + t \leqslant p,$$

and
$$\max_i j_i \leqslant p - r - t + 1.$$

To substantiate these claims, consider what happens if Γ is differentiated once with respect to s. We obtain a number of terms:

(a) one by increasing t by 1;

(b) r terms by increasing each j_i by 1, in turn;

(c) one term by increasing r by 1 in the denominator and adding a factor ${}^*F^{(1)}$ to the numerator.

It should now be evident that an inductive argument will prove the correctness of our assertions concerning Γ.

We shall next rewrite Γ, using Lemmas 4, 6 and 13, in the form

$$\Gamma = \gamma_p {}^*F^{(j_1)}{}^*F^{(j_2)}\ldots{}^*F^{(j_r)}\{{}^*F_{(1)}\}^t {}_*\phi_{n+r+t}.$$

This equation can be written, more simply, as

$$\Gamma = \gamma_p {}^*G(s) {}_*\phi_{n+r+t}, \tag{4·13}$$

where G is a distribution function whose first

$$(n+p+1)-(p-r-t+1) = n+r+t$$

moments are finite and representable as γ_{n+p+1}. A consequence of this is that, whenever $\rho \leqslant n+r+t$, we can use Lemma 4 to infer that

$$\frac{{}^*G(s)}{s^{\rho+1}} = \frac{1}{s^{\rho+1}} + \frac{\gamma_{n+p+1}}{s^{\rho}} + \ldots + \frac{\gamma_{n+p+1}}{s^2} + \gamma_{n+p+1} {}_*G_{[\rho]}(s).$$

If we couple this last observation with (4·9), in which n is to be replaced by $n+r+t$, then we can speedily deduce from (4·13) the representation

$$\Gamma = \frac{\gamma_{n+p+1}}{s^{n+r+t+1}} + \frac{\gamma_{n+p+1}}{s^{n+r+t}} + \ldots + \frac{\gamma_{n+p+1}}{s^2} + {}_*\Omega,$$

where ${}_*\Omega$ is to represent the L.T. of a function of B.V. But Γ is a typical term in the expression for $(-1)^p {}_*\phi_n^{(p)}(s)$, and hence we can infer that

$$(-1)^p {}_*\phi_n^{(p)}(s) = \frac{\gamma_{n+p+1}}{s^{n+p+1}} + \frac{\gamma_{n+p+1}}{s^{n+p}} + \ldots + \frac{\gamma_{n+p+1}}{s^2} + {}_*\Omega, \tag{4·14}$$

where ${}_*\Omega$ continues to represent the L.T. of some (different) function of B.V.

At this stage we invoke the preliminary result (4·10) to infer that many of the coefficients in the expansion (4·14) must vanish identically, and that

$$(-1)^p {}_*\phi_n^{(p)}(s) = \frac{\gamma_1}{s^{n+p+1}} + \frac{\gamma_2}{s^{n+p}} + \ldots + \frac{\gamma_{n+1}}{s^{p+1}} + {}_*\Omega.$$

Thus
$$\phi_n(t) = \gamma_1 t^n + \gamma_2 t^{n-1} + \ldots + \gamma_{n+1} + \frac{\Omega(t)}{t^p}. \tag{4·15}$$

Moreover, since $\Omega(t)$ is B.V., it tends to a limit $\Omega(+\infty)$ as $t \to +\infty$. Thus $s\,{}_*\Omega(s) \to \Omega(+\infty)$ as $s \to 0+$. But we must have $s\,{}_*\Omega(s) = o(1)$, for (4·14) to agree with (4·10). Hence $\Omega(+\infty) = 0$; and since the remainder term $t^{-p}\Omega(t)$ must be bounded (by Theorem 1), we can rewrite this remainder term as $\lambda(t)/(1+t)^p$, where $\lambda(t)$ is of B.V. and vanishes at $+\infty$.

We shall next prove: (*a*) that $\lambda(t)$ satisfies (*) and hence belongs to $\mathscr{B}$; (*b*) that $\lambda(t)$ belongs to $\mathscr{L}$, and hence to $\mathscr{R}$. It will plainly be enough to prove (*a*) and (*b*) for $\Omega(t)$ instead of $\lambda(t)$. Let us, temporarily, assume that we are given only that $\mu_{n+p} < \infty$ and not the full information $\mu_{n+p+1} < \infty$. Then the argument that led to (4·14) is still possible, but we can no longer assume Ω to be B.V. However, reference to the argument following (4·13) shows that ${}_*\Omega$ is the sum of terms of the forms

(i) $\gamma_{n+p+1} {}_*G_{[\rho]}$ $(\rho \leqslant n+r+t)$,

(ii) ${}_*G * M_{n+r+t}^{n+r+t}$ $(r+t \leqslant p)$.

All terms of type (i) are L.T.'s of functions of B.V. and must automatically satisfy (**). By Lemma 11 (ii) the terms of type (ii) are also L.T.'s of functions which satisfy (**). Thus Ω satisfies (**). In other words we have shown that *if* $\mu_{n+p+1} < \infty$ *then we may write*

$$\phi_n(t) = \gamma_1 t^n + \gamma_2 t^{n-1} + \ldots + \gamma_{n+1} + \frac{\Lambda(t)}{t^{1+p}}, \tag{4·16}$$

where $\Lambda(t)$ satisfies (**). If we compare (4·15) with (4·16) we deduce that $\Omega(t) = \Lambda(t)/t$. Hence

$$\Omega(t)-\Omega(t-\alpha) = \frac{\Lambda(t)-\Lambda(t-\alpha)}{t} - \frac{\alpha}{t}\Omega(t-\alpha),$$

and the demonstration that $\Omega(t)$ satisfies (*) follows from the facts that $\Lambda(t)$ satisfies (**) and that $\Omega(t)$ vanishes at $+\infty$. This proves (a).

To prove (b) let us suppose temporarily that we are given that $\mu_{n+p+2} < \infty$. Consider terms of type (i). The distribution function G involved in a typical term was shown above to have its first $n+r+t$ moments finite, when $\mu_{n+p+1} < \infty$. We are now assuming that $\mu_{n+p+2} < \infty$, and so we can take it that G has its first $n+r+t+1$ moments finite. Hence, the first moment of $G_{[\rho]}$ will be finite, and representable as γ_{n+p+2}. By Lemma 4 we can therefore infer that terms of type (i) can always be written as

$$\frac{\gamma_{n+p+1}}{s} + \gamma_{n+p+2} {}^*G_{[p+1]}(s). \tag{4·17}$$

In addition, terms of type (ii) can, by Lemma 10 and the finiteness of μ_{n+p+2}, be written as

$$\frac{\gamma_{n+p+2}\,{}^*G(s)}{s} + {}^*G(s)\,{}^*M^{n+p}_{n+p+1}(s),$$

and thence as

$$\frac{\gamma_{n+p+2}}{s} + \gamma_{n+p+2}\,{}^*G_{[1]}(s) + {}^*G(s)\,{}^*M^{n+p}_{n+p+1}(s). \tag{4·18}$$

If we combine the new expressions (4·17) and (4·18) of all terms of types (i) and (ii) then it is evident that ${}_*\Omega$ can be expressed in the form

$$\frac{\gamma_{n+p+2}}{s} + {}^*\Xi(s),$$

where $\Xi(t)$ is some function of B.V. But a (necessary and) sufficient condition for a function to belong to $L_1(0,\infty)$ is that its L.T. be recognizable as the L.-S.T. of a function of B.V. Hence $\Omega(t) - \gamma_{n+p+2}$ belongs to L_1. But $\Omega(t)$ vanishes at $+\infty$. Thus $\Omega(t)$ belongs to L_1. What we have hereby succeeded in proving may now be stated as follows: *if* $\mu_{n+p+1} < \infty$ *then*

$$\phi_n(t) = \gamma_1 t^n + \gamma_2 t^{n-1} + \ldots + \gamma_{n+1} + \frac{\Pi(t)}{t^{p-1}}, \tag{4·19}$$

where $\Pi(t) \in L_1$. If we compare (4·19) with (4·15) we deduce that $\Pi(t) = \Omega(t)/t$. From this equation the proof of (b) is apparent. This brings to an end the proof of Theorem 2.

We have now proved the required polynomial approximations to the ϕ-moments, and are ready to turn, in the next section, to the study of the ψ-cumulants. Before proceeding, however, let us see what can be said about $\phi_n(t)$ if $\mu_{n+1} = \infty$. In particular, suppose we are only given that $\mu_{q+1} < \infty$, $q < n$. Then, as in the proof of Theorem 1, we can at least get as far as

$${}^*\phi_n(s) = \frac{\gamma_1}{s^n} + \frac{\gamma_2}{s^{n-1}} + \ldots + \frac{\gamma_q}{s^{n-q+1}} + \frac{{}^*M^q_q}{s^{n-q}} + \ldots + {}^*M^n_q.$$

The following is then an easy deduction from Lemma 12.

COROLLARY 1. *If* $\mu_{q+1} < \infty$, $q > 0$, *then*

$$\phi_n(t) = \gamma_1 t^n + \gamma_2 t^{n-1} + \ldots + \gamma_{q+1} t^{n-q} + (1+t)^{n-q}\lambda(t),$$

where $\lambda(t) \in \mathscr{B}$.

Our results would have greater symmetry if we were able to replace $\mathscr{B}$ by $\mathscr{R}$ in Corollary 1, but we have been unable to prove that this replacement is allowable.

5. Asymptotic behaviour of the ψ-cumulants

Theorem 3. *If* $\mu_{n+p+1} < \infty$, $p > 0$, *then*

$$\psi_n(t) = \gamma_n t + \gamma_{n+1} + \frac{\lambda(t)}{(1+t)^p},$$

where $\lambda(t) \in \mathscr{R}$.

Proof. $\psi_n(t)$ is a linear combination, with numerical weights, of a finite number of terms like

$$\phi_{p_1}(t)\,\phi_{p_2}(t) \dots \phi_{p_r}(t),$$

where $p_1 + p_2 + \dots + p_r = n$. By Theorem 2 we have

$$\phi_{p_j}(t) = \gamma_1 t^{p_j} + \gamma_2 t^{p_j - 1} + \dots + \gamma_{p_j+1} + \frac{\lambda_{p_j}(t)}{(1+t)^{n+p-p_j}},$$

where $\lambda_{p_j}(t) \in \mathscr{R}$. Hence

$$\phi_{p_1}(t)\,\phi_{p_2}(t) \dots \phi_{p_r}(t) = \gamma_1 t^n + \gamma_2 t^{n-1} + \dots + \gamma_n + R(t),$$

where $R(t)$ is the sum of a finite number of terms of the form

$$t^m \times \frac{\lambda_{l_1}(t)}{(1+t)^{n+p-l_1}} \times \frac{\lambda_{l_2}(t)}{(1+t)^{n+p-l_2}} \times \dots \times \frac{\lambda_{l_s}(t)}{(1+t)^{n+p-l_s}},$$

with $m + l_1 + l_2 + \dots + l_s \leqslant n$. By property (C 5) of the class $\mathscr{R}$, this form can be rewritten

$$\frac{\lambda(t)}{(1+t)^p} \left\{ \frac{t^m}{(1+t)^{sn+(s-1)p-\Sigma l_i}} \right\},$$

where $\lambda(t) \in \mathscr{R}$. But we must have $sn + (s-1)\,p - \Sigma\, l_i \geqslant m$, so the function in the braces is clearly in the class $\mathscr{B}$. A second appeal to (C 5) allows us to write

$$R(t) = \frac{\lambda(t)}{(1+t)^p}$$

for some $\lambda(t) \in \mathscr{R}$. Combining all such remainder terms we discover that

$$\psi_n(t) = \gamma_1 t^n + \gamma_2 t^{n-1} + \dots + \gamma_{n+1} + \frac{\lambda(t)}{(1+t)^p} \quad (\lambda \in \mathscr{R}). \tag{5·1}$$

The theorem will be proved if we can show that the coefficients $\gamma_1, \gamma_2, \dots, \gamma_{n-1}$ in the right-hand side of (5·1) all vanish. These coefficients depend on the first $n-1$ moments $\mu_1, \mu_2, \dots, \mu_{n-1}$ of $F(x)$ and on no other moments. It might be possible to follow through the algebra of their derivation and then to show that they automatically vanish, but the algebra involved seems prohibitive (although we return to this topic in the appendix). For the present proof we adopt a different attack.

Let $\mathscr{F}_{n-1}$ be the space of all distribution functions of non-negative random variables whose first $(n-1)$ moments are finite, and let $\mathscr{L}_{n-1}$ be the Euclidean space of $(n-1)$ dimensions. Then we can map $\mathscr{F}_{n-1}$ into $\mathscr{L}_{n-1}$ by letting $F \in \mathscr{F}_{n-1}$ map into the point $\mu = (\mu_1, \mu_2, \dots, \mu_{n-1})$ whose co-ordinates are the first $(n-1)$ moments, in the correct order, of F. Call this mapping T_{n-1} so that $T_{n-1} F = \mu$. Then we wish to show that the rational functions $\gamma_1, \gamma_2, \dots, \gamma_{n-1}$ of $(\mu_1, \mu_2, \dots, \mu_{n-1})$ vanish throughout $T_{n-1}\mathscr{F}_{n-1}$. It will be sufficient if we can find an $(n-1)$-sphere $S_{(n-1)} \subset T_{n-1}\mathscr{F}_{n-1}$ and show that these rational functions vanish throughout $S_{(n-1)}$. Let $\mathscr{G}_{n-1}$ be the class of $F(x)$ which can be written

$$1 - \frac{1}{(n-1)} \{e^{-x/\nu_1} + e^{-x/\nu_2} + \dots + e^{-x/\nu_{n-1}}\} \quad (x \geqslant 0),$$

where $\nu_1, \nu_2, \ldots, \nu_{n-1}$ are real and positive. We show below that $T_{n-1}\mathscr{G}_{n-1}$ contains an $S_{(n-1)}$. Thus it will, in fact, be sufficient to prove that $\gamma_1, \gamma_2, \ldots, \gamma_{n-1}$ vanish whenever $F \in \mathscr{G}_{n-1}$.

To prove that $T_{n-1}\mathscr{G}_{n-1}$ contains an $S_{(n-1)}$ we remark that if $F \in \mathscr{G}_{n-1}$ then the $\{\mu_r\}$ are given as elementary functions of $\nu_1, \nu_2, \ldots, \nu_{n-1}$:

$$\mu_r = \frac{r!}{(n-1)}\{\nu_1^r + \nu_2^r + \ldots + \nu_{n-1}^r\} \quad (r = 1, 2, \ldots).$$

It is simple to show that there is a point $(\nu_1, \nu_2, \ldots, \nu_{n-1})$ in a spherical neighbourhood of which the Jacobian

$$\frac{\partial(\mu_1, \mu_2, \ldots, \mu_{n-1})}{\partial(\nu_1, \nu_2, \ldots, \nu_{n-1})}$$

does not vanish. Thus there is a point $(\mu_1, \mu_2, \ldots, \mu_{n-1})$, a spherical neighbourhood of which is in $T_{n-1}\mathscr{G}_{n-1}$.

Suppose $F \in \mathscr{G}_{n-1}$ and generates a renewal process $X_1, X_2, \ldots$, in the usual way. Let us suppose $\nu_1 \leqslant \nu_i$, all i, and write

$$F^*(x) = 1 - e^{-x/\nu_1}.$$

Then we can uniquely determine a renewal process $X_1^*, X_2^*, \ldots$ by demanding $F^*(X_i^*) = F(X_i)$, all i. By the construction of F^* we see that $X_i^* \leqslant X_i$, all i, and hence, if we employ obvious extensions of our notation, $N_t^* \geqslant N_t$. This last inequality implies $\phi_n^*(t) \geqslant \phi_n(t)$ for all n, and so the generating function $\Phi_t(\zeta)$ is undoubtedly analytic in any circle $C_\epsilon \equiv \{\zeta: |\zeta| < \epsilon\}$ within which $\Phi_t^*(\zeta)$ is analytic. But a routine calculation shows that

$$\Phi_t^*(\zeta) = \frac{\exp\{\zeta t/\nu_1(1-\zeta)\}}{(1-\zeta)},$$

which is analytic in C_1. Thus $\Phi_t(\zeta)$ is analytic in C_1. We use this fact presently.

If $F \in \mathscr{G}_{n-1}$ then

$$*F(s) = \frac{1}{n-1}\sum_1^{n-1}\frac{1}{1+\nu_i s}. \tag{5·2}$$

Let $s = \sigma + i\tau$ with σ, τ real. Then it is evident from (5·2) that $*F(\sigma)$ is real and strictly decreasing in $I \equiv [0, +\infty)$, and that if $\sigma \epsilon I$ and $\tau \neq 0$ then $|*F(s)| < *F(\sigma)$. Thus the $(n-1)$ zeros of $*F(s) - 1$ must consist of one zero at $s = 0$ and $(n-2)$ zeros in the open half-plane $\sigma < 0$. Furthermore, all these $(n-1)$ zeros must be located at regular points of $*F(s)$, although some may be located at zeros of $*F'(s)$ so that the possibility of multiple zeros cannot be ruled out. Now consider the equation

$$*F(s) = 1 - \zeta, \tag{5·3}$$

where ζ is complex, and confined to some small neighbourhood, C_ϵ, of the origin. Then it follows from the previous discussion of this paragraph, and from the theory of implicit functions (Goursat (1933), vol. II, p. 287) that (5·3) has $(n-1)$ roots $z_1(\zeta), z_2(\zeta), \ldots, z_{n-1}(\zeta)$ which are continuous functions of ζ. With no loss of generality we can take it that $z_1(0) = 0$. Then, if ϵ, δ are sufficiently small, we shall have $z_2(\zeta), z_3(\zeta), \ldots, z_{n-1}(\zeta)$ confined to the open half-plane $\sigma < -2\delta$, $|z_1(\zeta)| < \delta$, whenever $\zeta \epsilon C_\epsilon$. However, because of the possibility of multiple zeros of $*F(s) - 1$ we cannot be sure that $z_2(\zeta), z_3(\zeta), \ldots, z_{n-1}(\zeta)$ always have distinct values unless we restrict ζ to an annulus $A_\epsilon \equiv \{\zeta: \frac{1}{2}\epsilon < |\zeta| < \epsilon\}$.

It will be convenient to write $*F(s) = Q(s)/P(s)$, where $Q(s)$ is a polynomial of degree $(n-2)$, $P(s)$ a polynomial of degree $(n-1)$. This is legitimate, by (5·2). If we write

$$*\Phi_s(\zeta) = \int_0^\infty e^{-st}\Phi_t(\zeta)\,dt,$$

then Lemma 6 leads to the conclusion

$$
\begin{aligned}
{}_*\Phi_s(\zeta) &= \frac{1}{s}[1-\zeta\{1-{}_*F(s)\}^{-1}]^{-1} \\
&= \frac{1}{s}\frac{P(s)-Q(s)}{(1-\zeta)P(s)-Q(s)}.
\end{aligned} \tag{5·4}
$$

Apart from the factor s, the polynomial in the denominator of (5·4) evidently vanishes at $z_1(\zeta), z_2(\zeta), \ldots, z_{n-1}(\zeta)$. Bearing in mind that $z_1(0) = 0$, some elementary manipulation shows that (5·4) may be replaced by

$$
{}_*\Phi_s(\zeta) = -\frac{\prod_{i=2}^{n-1}[s-z_i(0)]}{\prod_{i=1}^{n-1}[s-z_i(\zeta)]}. \tag{5·5}
$$

Thus we may adopt a partial fraction expansion

$$
{}_*\Phi_s(\zeta) = \sum_{i=1}^{n-1}\frac{Z_i(\zeta)}{s-z_i(\zeta)}, \tag{5·6}
$$

where

$$
Z_j(\zeta) = \frac{\prod_{i=2}^{n-1}[z_j(\zeta)-z_i(0)]}{\prod_{\substack{i=1\\ i\neq j}}^{n-1}[z_j(\zeta)-z_i(\zeta)]} \quad (j = 1, 2, \ldots, n-1). \tag{5·7}
$$

Since, for $\zeta \in A_\varepsilon$, the $z_i(\zeta)$ are distinct-valued continuous functions of ζ it follows from (5·7) that the $Z_j(\zeta)$ are finite-valued, continuous functions of $\zeta \in A_\varepsilon$. Thus we can find a bound $M < \infty$ such that $|z_j(\zeta)| < M$ for all j, uniformly in $\zeta \in A_\varepsilon$. Moreover, it is apparent that $Z_1(\zeta)$ cannot vanish in A_ε for this would require that one of the factors in the numerator of (5·7) vanishes in A_ε. But if $\zeta \in A_\varepsilon$ then $|z_1(\zeta)| < \delta$, and $|z_j(0)| > 2\delta$ for all $j \neq 1$. Thus $Z_1(\zeta)$ cannot vanish in A_ε and so we can find a bound $m > 0$ such that $|Z_1(\zeta)| > m$ for all $\zeta \in A_\varepsilon$.

We next invert the L.T. (5·6) and obtain

$$
\Phi_t(\zeta) = \sum_{i=1}^{n-1} Z_i(\zeta)\, e^{tz_i(\zeta)}. \tag{5·8}
$$

If we suppose $\zeta \in A_\varepsilon$ and use the results of the preceding paragraph then (5·8) can be rewritten

$$
\Phi_t(\zeta) = Z_1(\zeta)\, e^{tz_1(\zeta)}\{1 + S_1(\zeta, t)\},
$$

where $S_1(\zeta, t) = 0(e^{-\delta t})$ uniformly in ζ. If we take logarithms we have

$$
\Psi_t(\zeta) = tz_1(\zeta) + \log Z_1(\zeta) + S_2(\zeta, t), \tag{5·9}
$$

where $S_2(\zeta, t)$ is also $0(e^{-\delta t})$ uniformly in $\zeta \in A_\varepsilon$.

We have shown earlier that $\Phi_t(\zeta)$ is analytic in C_1 and it is plain from the definition of $\Phi_t(\zeta)$ that $\Phi_t(0) = 1$. Thus, in some C_ε, $\Psi_t(\zeta)$ is an analytic function of ζ. Hence

$$
\psi_n(t) = \frac{n!}{2\pi i}\oint_{C'}\frac{\Psi_t(\zeta)}{\zeta^{n+1}}\,d\zeta,
$$

where $C' = C_{\frac{1}{2}\varepsilon}$. If we now make use of (5·9) we can obtain the result

$$
\psi_n(t) = \alpha_n t + \beta_n + 0(e^{-\delta t}), \tag{5·10}
$$

where
$$\alpha_n = \frac{n!}{2\pi i}\oint_{C'} \frac{z_1(\zeta)}{\zeta^{n+1}}\,d\zeta, \tag{5·11}$$

$$\beta_n = \frac{n!}{2\pi i}\oint_{C'} \frac{\log Z_1(\zeta)}{\zeta^{n+1}}\,d\zeta. \tag{5·12}$$

A comparison of (5·1) and (5·10) shows that the rational functions $\gamma_1, \gamma_2, \ldots, \gamma_{n-1}$ undoubtedly vanish when $F \in \mathscr{G}_{n-1}$. This proves the theorem.

THEOREM 4. *If* $\mu_{n+1} < \infty$ *then*
$$\psi_n(t) = \gamma_n t + \gamma_{n+1} + \lambda(t),$$
where $\lambda(t) \in \mathscr{B}$.

Proof. This theorem is proved by the same argument as Theorem 4, with the difference that we must now appeal to Corollary 1 to deal with $\phi_n(t)$.

One can also deduce from Corollary 1 the following

COROLLARY 2. *If* $\mu_n < \infty$ *then*
$$\psi_n(t) = \gamma_n t + (1+t)\,\lambda(t),$$
where $\lambda(t) \in \mathscr{B}$.

6. THE FORMAL CALCULATIONS

Our position is that we have shown that under certain conditions $\psi_n(t) = \alpha_n t + \beta_n + r_n(t)$, where α_n is a rational function of $\mu_1, \mu_2, \ldots, \mu_n$; β_n is a rational function of $\mu_1, \mu_2, \ldots, \mu_{n+1}$, and $r_n(t)$ is a remainder term tending to zero at $+\infty$. We now turn our attention to the formal calculation of the rational functions $\alpha_n(\mu_1, \mu_2, \ldots, \mu_n), \beta_n(\mu_1, \mu_2, \ldots, \mu_{n+1})$.

To justify the method we adopt, consider the problem of calculating $\beta_n(\mu_1, \mu_2, \ldots, \mu_{n+1})$ defined on $\mathscr{L}_{n+1}$. If two rational functions on $\mathscr{L}_{n+1}$ are identically equal throughout an $(n+1)$-sphere $S_{(n+1)} \subset \mathscr{L}_{n+1}$ then they must be identically equal throughout $\mathscr{L}_{n+1}$. But $T_{n+1}\mathscr{G}_{n+1}$ contains an $S_{(n+1)}$. Thus we may suppose, in the calculation of $\beta_n(\mu_1, \mu_2, \ldots, \mu_{n+1})$ that $F \in \mathscr{G}_{n+1}$. In fact we shall proceed as though F is an arbitrary member of $\mathscr{G} = \bigcup_n \mathscr{G}_n$; if we calculate α_n and β_n accordingly then our results must be correct for all values of n.

If $F \in \mathscr{G}$ then: (i) $*F(s)$ is analytic in some open neighbourhood of $s = 0$; (ii) there is a unique analytic function $z_1(s)$ defined on this neighbourhood, vanishing at $s = 0$, and satisfying the relation $s - 1 = *F(z_1(s))$. This all flows from the discussion in § 5. Moreover, by Cauchy's integral and (5·11) we can deduce that if $|s| < \frac{3}{4}\epsilon$ then

$$\begin{aligned} z_1(s) &= \frac{1}{2\pi i}\oint_{C'} \frac{z_1(\zeta)}{\zeta - s}\,d\zeta, \\ &= \frac{1}{2\pi i}\oint_{C'} \frac{z_1(\zeta)}{\zeta}\left\{\sum_0^\infty \left(\frac{s}{\zeta}\right)^n\right\} d\zeta, \\ &= \sum_0^\infty \frac{\alpha_n}{n!}\,s^n. \end{aligned} \tag{6·1}$$

Uniform convergence justifies the change of position of the integration and summation signs. The equation (6·1) suggests that one way of calculating the $\{\alpha_n\}$ is to expand $z_1(s)$ as a Taylor series at the origin. But we have

$$s = \sum_{n=1}^\infty \frac{(-1)^{n+1}\mu_n}{n!}\,z_1^n(s)$$

by the definition of $z_1(s)$; in other words we have s as a Taylor series in z_1.† Our problem is therefore merely one of inverting a series. This process has been discussed by many authors, and the fullest computations appear to be those given by Van Orstrand (1910), but one may also refer to Adams & Hippisley(1922), Dwight (1957), and Goursat (1933, vol. I, p. 474). If we use Van Orstrand's results we can obtain the values of α_n listed below. In point of fact our values of α_n were calculated independently, by a method which we describe in §7 and depends on the correctness of the calculations for β_n, and then compared with Van Orstrand's results. This procedure provided an additional check on our β_n calculations. In quoting our results we take $\mu_1 = 1$. We found:

$$\alpha_2 = \mu_2,$$

$$\alpha_3 = -\mu_3 + 3\mu_2^2,$$

$$\alpha_4 = \mu_4 - 10\mu_2\mu_3 + 15\mu_2^3,$$

$$\alpha_5 = -\mu_5 + 15\mu_2\mu_4 + 10\mu_3^2 - 105\mu_2^2\mu_3 + 105\mu_2^4,$$

$$\alpha_6 = \mu_6 - 21\mu_2\mu_5 - 35\mu_3\mu_4 + 210\mu_2^2\mu_4 + 280\mu_2\mu_3^2 - 1260\mu_2^3\mu_3 + 945\mu_2^5,$$

$$\alpha_7 = -\mu_7 + 28\mu_2\mu_6 + 56\mu_3\mu_5 - 378\mu_2^2\mu_5 + 35\mu_4^2 - 1260\mu_2\mu_3\mu_4 + 3150\mu_2^3\mu_4 - 280\mu_3^3 + 6300\mu_2^2\mu_3^2 - 17{,}325\mu_2^4\mu_3 + 10{,}395\mu_2^6,$$

$$\alpha_8 = \mu_8 - 36\mu_2\mu_7 - 84\mu_3\mu_6 + 630\mu_2^2\mu_6 - 126\mu_4\mu_5 + 2520\mu_2\mu_3\mu_5 - 6930\mu_2^3\mu_5 + 1575\mu_2\mu_4^2 + 2100\mu_3^2\mu_4 - 34{,}650\mu_2^2\mu_3\mu_4 + 51{,}975\mu_2^4\mu_4 - 15{,}400\mu_2\mu_3^3 + 138{,}600\mu_2^3\mu_3^2 - 270{,}270\mu_2^5\mu_3 + 135{,}135\mu_2^7.$$

Let us now turn to the more involved problem of calculating β_n. Lemma 4 tells us there is a distribution function $F_{(1)}$ such that

$$*F_{(1)}(s) = \frac{1 - *F(s)}{\mu_1 s},$$

and if, for typographical simplicity, we write $\tau_1, \tau_2, \ldots$ for the moments of $F_{(1)}$, then

$$*F_{(1)}(s) = \sum_{r=0}^{\infty} \frac{\tau_r (-s)^r}{r!}. \tag{6·2}$$

By Lemma 6 and (6·2) we have therefore that

$$\frac{*\phi_n(s)}{n!} = \frac{1}{s^n}\left\{1 - \tau_1 s + \frac{\tau_2}{2!}s^2 - \ldots\right\}^{-n},$$

and the coefficient of s^0 in the Laurent expansion of the right-hand side of this equation is the coefficient of s^n in

$$1 + n\left[\tau_1 s - \frac{\tau_2}{2!}s^2 + \ldots\right] + \frac{n(n+1)}{2!}\left[\tau_1 s - \frac{\tau_2}{2!}s^2 + \ldots\right]^2 + \frac{n(n+1)(n+2)}{3!}\left[\tau_1 s - \frac{\tau_2}{2!}s^2 + \ldots\right]^3 + \text{etc.},$$

which is evidently

$$\sigma_n = (-1)^n \sum_{d=1}^{d=n} \left\{(-1)^d n(n+1)\ldots(n+d-1)\,\Sigma^* \frac{\tau_{\pi_1}^{\rho_1}\tau_{\pi_2}^{\rho_2}\ldots\tau_{\pi_\nu}^{\rho_\nu}}{\rho_1!\rho_2!\ldots\rho_\nu!(\pi_1!)^{\rho_1}(\pi_2!)^{\rho_2}\ldots(\pi_\nu!)^{\rho_\nu}}\right\}, \tag{6·3}$$

where the inner Σ^* is a summation over all positive integers $\pi_1, \pi_2, \ldots, \pi_\nu; \rho_1, \rho_2, \ldots, \rho_\nu$ such that the π_i are all different, $\Sigma\rho\pi = n$ and $\Sigma\rho = d$.

† For a heuristic discussion which shows the present significance of the root $z_1(s)$, see Bartlett (1949).

Suppose a function $Q(t)$, say, can be written as $P(t)+r(t)$, where $P(t)$ is a polynomial in t and $r(t)$ tends to zero, as t tends to $+\infty$, faster than any power of t. Let us agree to mean by 'the constant term in $Q(t)$' the coefficient of t^0 in $P(t)$ and by the 'linear coefficient in $Q(t)$' let us mean the coefficient of t in $P(t)$. When $F \in \mathscr{G}$ all the moments of F are finite and we can therefore speak of the constant term in $\phi_n(t)$, by Theorem 2. But the constant term in $\phi_n(t)$ must be the coefficient of s^0 in the Laurent expansion of ${}^*\phi_n(s)$. Thus the constant term in $\phi_n(t)/n!$ is σ_n, given in (6·3).

It is trivial to verify that the constant term in, say, $\phi_p(t)\,\phi_q(t)$ is merely the product of the constant term in $\phi_p(t)$ and the constant term in $\phi_q(t)$, and it is equally trivial to extend this result to products of several ϕ_n's and thence to linear combinations of products of several ϕ_n's. But, by the well-known relation between moments and cumulants (Kendall & Stuart (1958)),

$$\frac{\psi_n}{n!} = \sum_{d=1}^{d=n} \left\{ (-1)^{d-1}(d-1)!\,\Sigma^* \frac{\phi_{\pi_1}^{\rho_1}\phi_{\pi_2}^{\rho_2}\dots\phi_{\pi_\nu}^{\rho_\nu}}{\rho_1!\,\rho_2!\dots\rho_\nu!\,(\pi_1!)^{\rho_1}(\pi_2!)^{\rho_2}\dots(\pi_\nu!)^{\rho_\nu}} \right\}, \tag{6·4}$$

and so we deduce that β_n, the constant term in $\psi_n(t)$, is given by

$$\frac{\beta_n}{n!} = \sum_{d=1}^{d=n} \left\{ (-1)^{d-1}(d-1)!\,\Sigma^* \frac{\sigma_{\pi_1}^{\rho_1}\sigma_{\pi_2}^{\rho_2}\dots\sigma_{\pi_\nu}^{\rho_\nu}}{\rho_1!\,\rho_2!\dots\rho_\nu!} \right\}. \tag{6·5}$$

We shall make use of the customary terminology: *weight* and *degree*. For instance, the τ-product $\tau_1^2\tau_3\tau_5^3$ is of degree $2+1+3=6$ and weight $2\times1+1\times3+3\times5=20$. It is to be noticed from (6·3) that σ_n is a certain linear combination of all τ-products of weight n. Similarly, from (6·5), we see that β_n is a certain linear combination of all σ-products of weight n. Hence β_n is a certain linear combination of all τ-products of weight n, and our method of computing β_n will involve taking each τ-product of weight n, in turn, and determining its coefficient in β_n. (It is worth remarking that this is another place where it has been to our advantage to be dealing with ψ-cumulants instead of conventional ones. As can be seen, now, from (2·6) the constant term b_n in the conventional cumulant $k_n(t)$ involves all τ-products of weight $\leqslant n$.)

To be specific for a moment, consider a term $\tau_1\tau_2^2$ of weight 5 and so occurring in β_5. How can $\tau_1\tau_2^2$ arise from (6·5)? In various ways. We may obtain $\tau_1\tau_2^2$ in the expansions of σ_5, $\sigma_1\sigma_4$, $\sigma_2\sigma_3$, or $\sigma_1\sigma_2^2$. Thus, if we represent $\tau_1\tau_2^2$ by the multipartite number $(1, 2^2)$ the various ways in which $\tau_1\tau_2^2$ will arise in (6·5) correspond to the various separations $(1, 2^2)$, $(1)\,(2^2)$, $(2)\,(1, 2)$, $(1)\,(2)^2$. For each separation we must calculate the individual contribution to the coefficient of $\tau_1\tau_2^2$ in β_5.

Suppose we now consider a general separation

$$\Pi \equiv S_{11}^{r_{11}} S_{12}^{r_{13}} \dots S_{1a}^{r_{1a}} S_{21}^{r_{21}} \dots S_{2b}^{r_{2b}} \dots S_{kf}^{r_{kf}}$$

of some general multipartite number S. In this $S_{11}, S_{12}, \dots, S_{1a}$ are a different multipartite numbers of the same weight w_1; $S_{21}, S_{22}, \dots, S_{2b}$ are b different multipartite numbers of weight w_2; and so on. Then the weight of S is $w = w_1r_1 + w_2r_2 + \dots + w_kr_k$, where $r_1 = r_{11} + r_{12} + \dots + r_{1a}$, etc. This general separation represents one set of ways in which the τ-product T corresponding to S could arise in (6·5). In fact, the term in (6·5) which is involved is

$$\sigma_{w_1}^{r_1}\,\sigma_{w_2}^{r_2}\dots\sigma_{w_k}^{r_k}. \tag{6·6}$$

If we write $r = r_1 + r_2 + \dots + r_k$ then it is evident that the coefficient of (6·6) in (6·5) is

$$\frac{(-1)^{r-1}\,(r-1)!}{r_1!\,r_2!\dots r_k!}. \tag{6·7}$$

Now consider the term $\sigma_{w_1}^{r_1}$ which must give rise to a τ-product corresponding to the separation $S_{11}^{r_{11}} S_{12}^{r_{12}} \dots S_{1a}^{r_{1a}}$. By the multinomial theorem this τ-product can arise in

$$\frac{r_1!}{r_{11}!\, r_{12}! \dots r_{1a}!}$$

ways. Similarly for the τ-products arising from $\sigma_{w_2}^{r_2}, \dots, \sigma_{w_k}^{r_k}$. Thus the general separation Π arises from the term (6·6) in

$$\frac{r_1!\, r_2! \dots r_k!}{r_{11}!\, r_{12}! \dots r_{1a}!\, r_{21}!\, r_{22}! \dots r_{2b}! \dots r_{kf}!} \tag{6·8}$$

ways. If we combine the coefficient (6·7), which must multiply every τ-product from (6·6), with the number of ways (6·8) in which T may arise from (6·6) then we derive the *form factor*

$$\frac{(-1)^{r-1}(r-1)!}{r_{11}!\, r_{12}! \dots r_{kf}!}. \tag{6·9}$$

Notice that the form factor (6·9) depends only on the general 'shape' of the separation Π and not on the weights of the individual separates. Thus, for example, $(1, 2^2, 3)\,(2)^2\,(4)$ and $(1)\,(3)^2\,(7)$ have the same 'shape' which we will denote $(\cdot)\,(\cdot)\,(\cdot)^2$. We can thus construct a table of form factors which may be used for a variety of different separations.

The form factor (6·9) is merely the number of ways in which T can arise from (6·6), multiplied by the number (6·7) which must be a factor of every contribution from (6·6), to the final coefficient in β_w. Each separate S_{ij} in Π will also contribute a *numerical factor* which we now deduce from (6·3). Suppose $S_{ij} \equiv (\pi_1^{\rho_1}\, \pi_2^{\rho_2} \dots \pi_l^{\rho_l})$, corresponding to the τ-product

$$T_{ij} \equiv \tau_{\pi_1}^{\rho_1} \tau_{\pi_2}^{\rho_2} \dots \tau_{\pi_l}^{\rho_l}$$

with

$$w_i = \rho_1 \pi_1 + \rho_2 \pi_2 + \dots + \rho_l \pi_l,$$

$$d_{ij} = \rho_1 + \rho_2 + \dots + \rho_l.$$

Then the coefficient of T_{ij} in (6·3) is evidently

$$\frac{(-1)^{w_i + d_{ij}}\, w_i (w_i+1) \dots (w_i + d_{ij} - 1)}{\rho_1!\, \rho_2! \dots \rho_\nu!\, (\pi_1!)^{\rho_1} \dots (\pi_\nu!)^{\rho_\nu}}. \tag{6·10}$$

Let us write

$$c_{ij} = \frac{w_i (w_i+1) \dots (w_i + d_{ij} - 1)}{\rho_1!\, \rho_2! \dots \rho_\nu!}$$

for the *separate factor* of S_{ij}. Then the numerical factor (6·10) may be rewritten

$$(-)^{w_i + d_{ij}} \frac{c_{ij}}{(\pi_1!)^{\rho_1} \dots (\pi_\nu!)^{\rho_\nu}}. \tag{6·11}$$

Suppose now that

$$T \equiv \tau_{u_1}^{v_1} \tau_{u_2}^{v_2} \dots \tau_{u_l}^{v_l},$$

is of weight w and degree d. Then when we multiply the numerical factors (6·11) together, one for each separate S_{ij}, we plainly must end with

$$\left\{ \frac{(-)^{w+d}}{(u_1!)^{v_1}\, (u_2!)^{v_2} \dots (u_l!)^{v_l}} \right\} \prod_{i,j} c_{ij}. \tag{6·12}$$

In the resultant numerical multiplier (6·12) it may be seen that the factor in braces depends only on T and not on the particular separation Π involved. We may call this a *τ-factor*.

Thus to determine the coefficient of T in $\beta_w / w!$ we first determine its *τ-factor*. Then represent T by the appropriate multipartite number S and write down every possible separation

of S. For each separation we multiply together the *separate factors* of each separate, and multiply the product by the *form factor* of the separation. The resulting products for each separation of S are then summed, and the total multiplied by the τ-factor of T. The result is the coefficient of T in $\beta_w/w!$. To convert the result to an expression involving μ's instead of τ's we use the result

$$\tau_r = \frac{\mu_{r+1}}{(r+1)\mu_1}.$$

Example. Find the coefficient of $\tau_1\tau_2^2$ in $\beta_5/5!$.

(i) The τ-factor is evidently $\frac{(-)^{5+3}}{(2!)^2} = \frac{1}{4}$.

(ii) $S \equiv (1, 2^2)$ and the possible separations of S are

$$\Pi_1 \equiv (1, 2^2); \quad \Pi_3 \equiv (1, 2)(2);$$
$$\Pi_2 \equiv (1)(2^2); \quad \Pi_4 \equiv (1)(2)^2.$$

We need the following short tables.

Separate	Factor
(1)	1
(2)	2
(1, 2)	12
(2^2)	10
$(1, 2^2)$	105

Form	Factor
$(\cdot)$	+1
$(\cdot)(\cdot)$	−1
$(\cdot)(\cdot)^2$	+1

Using the tabulated separate and form factors we can compute the required coefficient as follows.

Separation	Form factor	Separation factor	Product
Π_1	+1	105	+105
Π_2	−1	1×10	− 10
Π_3	−1	12×2	− 24
Π_4	+1	1×2^2	+ 4
		Total	+ 75

If we multiply the total by the τ-factor $\frac{1}{4}$ we find that $\beta_5/5!$ involves a term $75\,\tau_1\tau_2^2/4$. It is an easy deduction from this result that β_5 involves a term

$$\frac{75 \times 5!}{4 \times 2 \times 3^2}\mu_2\mu_3^2 = 125\mu_2\mu_3^2,$$

(we have taken $\mu_1 = 1$).

For this example we have purposely chosen a fairly simple case, but the calculations become increasingly tedious as we deal with β's of higher weights. The chief source of error to be guarded against is that some separations of a multipartite number may be overlooked.

For instance, the multipartite number $(1^5, 2)$ referring to the term $\mu_2^5\mu_3$ in β_7 has 19 different separations.

By the method outlined above we have computed formulae for $\beta_1, \beta_2, \ldots, \beta_8$. We now list our results, taking $\mu_1 = 1$.

$$\beta_1 = \tfrac{1}{2}\mu_2,$$

$$\beta_2 = -\tfrac{2}{3}\mu_3 + \tfrac{5}{4}\mu_2^2,$$

$$\beta_3 = \tfrac{3}{4}\mu_4 - 5\mu_2\mu_3 + \tfrac{11}{2}\mu_2^3,$$

$$\beta_4 = -\tfrac{4}{5}\mu_5 + \tfrac{17}{2}\mu_2\mu_4 + \tfrac{16}{3}\mu_3^2 - 44\mu_2^2\mu_3 + \tfrac{279}{8}\mu_2^4,$$

$$\beta_5 = \tfrac{5}{6}\mu_6 - 13\mu_2\mu_5 - 20\mu_3\mu_4 + \tfrac{395}{4}\mu_2^2\mu_4 + 125\mu_2\mu_3^2 - 465\mu_2^3\mu_3 + \tfrac{579}{2}\mu_2^5,$$

$$\begin{aligned}\beta_6 = {} & -\tfrac{6}{7}\mu_7 + \tfrac{37}{2}\mu_2\mu_6 + 34\mu_3\mu_5 - 195\mu_2^2\mu_5 + \tfrac{165}{8}\mu_4^2 - 605\mu_2\mu_3\mu_4 + 1290\mu_2^3\mu_4 - \tfrac{1160}{9}\mu_3^3 \\ & + 2465\mu_2^2\mu_3^2 - 5790\mu_2^4\mu_3 + (11{,}895/4)\,\mu_2^6,\end{aligned}$$

$$\begin{aligned}\beta_7 = {} & \tfrac{7}{8}\mu_8 - 25\mu_2\mu_7 - \tfrac{161}{3}\mu_3\mu_6 + 350\mu_2^2\mu_6 - 77\mu_4\mu_5 + 1295\mu_2\mu_3\mu_5 - 3108\mu_2^3\mu_5 + \tfrac{1575}{2}\mu_2\mu_4^2 \\ & + 1015\mu_3^2\mu_4 - 14{,}560\mu_2^2\mu_3\mu_4 + (76{,}335/4)\,\mu_2^4\mu_4 - 6230\mu_2\mu_3^3 + 48{,}860\mu_2^3\mu_3^2 - 83{,}265\mu_2^5\mu_3 \\ & + (72{,}855/2)\,\mu_2^7,\end{aligned}$$

$$\begin{aligned}\beta_8 = {} & -\tfrac{8}{9}\mu_9 + \tfrac{65}{2}\mu_2\mu_8 + 80\mu_3\mu_7 - 584\mu_2^2\mu_7 + 133\mu_4\mu_6 - 2520\mu_2\mu_3\mu_6 + 6713\mu_2^3\mu_6 + \tfrac{392}{5}\mu_5^2 \\ & - 3626\mu_2\mu_4\mu_5 - 2352\mu_2^2\mu_5 + 37{,}464\mu_2^2\mu_3\mu_5 - 54{,}138\mu_2^4\mu_5 - 2870\mu_3\mu_4^2 + (91{,}315/4)\,\mu_2^2\mu_4^2 \\ & + 59{,}010\mu_2\mu_3^2\mu_4 - 339{,}990\mu_2^3\mu_3\mu_4 + (634{,}935/2)\,\mu_2^5\mu_4 + (19{,}040/3)\,\mu_3^4 - 218{,}960\mu_2^2\mu_3^3 \\ & + 1{,}019{,}970\mu_2^4\mu_3^2 - 1{,}359{,}960\mu_2^6\mu_3 + (8{,}294{,}895/16)\mu_2^8.\end{aligned}$$

To conclude this section let us remark, what should be fairly obvious, that one can obtain the $\{a_n\}$ and $\{b_n\}$ of Theorem 0, which refers to conventional cumulants, from the $\{\alpha_n\}$ and $\{\beta_n\}$ by means of the relations (2·6) and (2·7).

7. Checks on the formal calculations

We begin by describing the way in which we obtained the $\{\alpha_n\}$ from the $\{\beta_n\}$ calculations. This method is not the best way of determining the $\{\alpha_n\}$ directly; that developed in § 6, using Van Orstrand's results, is undoubtedly better. But if one has already performed the calculations of $\{\beta_n\}$ then the present method of deriving the $\{\alpha_n\}$ is fairly easy. By comparing these results with those obtainable by the method of § 6 we have a good check that no separations have been overlooked in the $\{\beta_n\}$ computations.

Let ω_n be the linear coefficient in $\phi_n(t)/n!$. Then ω_n will be the coefficient of s^{-1} in the Laurent expansion of $^*\phi_n(s)/n!$. By the same expansion which preceded (6·3) we have therefore that

$$\omega_1 = 1,$$

$$\omega_{n+1} = (-1)^n \sum_{d=1}^{d=n} \left\{ (-1)^d (n+1)(n+2)\ldots(n+d)\, \Sigma^* \frac{\tau_{\pi_1}^{\rho_1} \tau_{\pi_2}^{\rho_2} \ldots \tau_{\pi_\nu}^{\rho_\nu}}{\rho_1! \ldots \rho_\nu! \, (\pi_1!)^{\rho_1} \ldots (\pi_\nu!)^{\rho_\nu}} \right\}. \tag{7·1}$$

Thus ω_{n+1} is a linear combination of all τ-products of weight n. Notice, furthermore, that an individual τ-product of degree d in (7·1) has the same coefficient as it would have in σ_n, the constant term of $\phi_n(t)/n!$, except for being multiplied by a 'correction factor' $(n+d)/n$.

It will be easier to understand the argument that follows if we talk about a specific τ-product, so let us consider $\tau_1\tau_2^2$ (which occurs in $\alpha_6/6!$). Consider, in particular, the separation $(1, 2)(2)$ of the associated multipartite number $(1, 2^2)$. Corresponding to $(1, 2)(2)$ we

have a number of ways in which $\tau_1\tau_2^2$ can arise in (6·4) (with $n = 6$) in the linear coefficient $\alpha_6/6!$. These ways are:

(i) $\tau_1\tau_2$ from ω_4 in $\phi_4(t)/4!$, and τ_2 from σ_2 in $\phi_2(t)/2!$;

(ii) $\tau_1\tau_2$ from σ_3 in $\phi_3(t)/3!$, and τ_2 from ω_3 in $\phi_3(t)/3!$;

(iii) $\tau_1\tau_2$ from σ_3 in $\phi_3(t)/3!$, τ_2 from σ_2 in $\phi_2(t)/2!$, and the linear coefficient $\omega_1 = 1$ from $\phi_1(t)$.

Let $\mathscr{D}$ be the coefficient of $\tau_1\tau_2^2$ in $\alpha_6/6!$ and $\mathscr{C}$ the contribution of the separation (1, 2) (2) towards the coefficient of $\tau_1\tau_2^2$ in $\beta_5/5!$, as described in § 6. Then a careful investigation of (6·4) shows that (i) contributes an amount $(1+\frac{2}{3})\mathscr{C}$ towards $\mathscr{D}$. The factor $(1+\frac{2}{3})$ arises as the 'correction factor' of (1, 2). Similarly (ii) contributes an amount $(1+\frac{1}{2})\mathscr{C}$ towards $\mathscr{D}$.

Suppose Δ is a given 'shape' of p parts (e.g. $(\cdot)(\cdot)^2(\cdot)^3$ and $(\cdot)(\cdot)(\cdot)(\cdot)(\cdot)^2$ are 'shapes' of six parts). Let Δ have a form factor f. Then the augmented 'shape' $(\cdot)\Delta$ has form factor $-pf$. This contention is easily checked by (6·9).

Now consider the contribution from (iii) towards $\mathscr{D}$. Since both $\tau_1\tau_2$ and τ_2 are here behaving as from constant terms (σ's) no correction factor is required for them. But $\mathscr{C}$ is based on a form factor of $(\cdot)(\cdot)$, whereas what is required for (iii) is a form factor of $(\cdot)(\cdot)(\cdot)$. By the preceding paragraph we see that (iii) contributes an amount $-2\mathscr{C}$ to $\mathscr{D}$. Hence the total contribution to $\mathscr{D}$ by the separation (1, 2) (2) is $(\frac{2}{3}+\frac{1}{2})\mathscr{C}$.

The preceding discussion gives an indication of the general result which we now describe, and which can be proved by a careful extension of the argument. Let T be a τ-product of weight w, and associated multipartite number S. Let $S_1^{r_1}S_2^{r_2}\dots S_k^{r_k}$ be a separation of S, where the separate S_j has weight w_j and degree d_j. Let $\mathscr{C}$ be the contribution from the separation $S_1^{r_1}S_2^{r_2}\dots S_k^{r_k}$ to the coefficient of T in $\beta_w/w!$. Then the contribution of this separation to the coefficient of T in $\alpha_{w+1}/(w+1)!$ is

$$\left\{\frac{r_1d_1}{w_1}+\frac{r_2d_2}{w_2}+\dots+\frac{r_kd_k}{w_k}\right\}\mathscr{C}. \tag{7·2}$$

It is this result that enables us to obtain the $\{\alpha_n\}$ rapidly from the $\{\beta_n\}$ calculations.

Example. Find the coefficient of $\tau_1\tau_2^2$ in $\alpha_6/6!$. We borrow the results of the example worked in § 6. In the computation scheme given below, column A is the extreme right-hand column of the computation scheme displayed in the example in § 6; in other words it consists of the various contributions to the coefficient of $\tau_1\tau_2^2$ in $\beta_5/5!$, but with the τ-factor left out. In the column B below, we give the appropriate value of the expression in braces in (7·2).

Separation	A	B	Product AB
$(1, 2^2)$	$+105$	$\frac{3}{6}$	$+63$
$(1)\,(2^2)$	$-\ 10$	$1+\frac{2}{4}$	-15
$(1, 2)\,(2)$	$-\ 24$	$\frac{2}{3}+\frac{1}{2}$	-28
$(1)\,(2)^2$	$+\ 4$	$1+\frac{2}{2}$	$+\ 8$
		Total	$+28$

The τ-factor, as before, is $\frac{1}{4}$.

Thus the coefficient of $\tau_1\tau_2^2$ in $\alpha_6/6!$ is $\frac{1}{4}\times 28 = 7$.

Converting to μ's instead of τ's, and taking $\mu_1 = 1$, we see that α_6 contains a term

$$\frac{7\times 6!}{2\times 3^2}\mu_2\mu_3^2 = 280\,\mu_2\mu_3^2.$$

This result will be found to agree with the value of α_6 given in § 6.

The most obvious check to employ is obtained by putting $F(x) = 1-e^{-x}$, so that $\mu_n = n!$ for all n. Routine computation shows that in this case $\psi_n(t) = (n!)\,t+(n-1)!$; thus the substitution $\mu_n = n!$ in our formulae for α_n and β_n should give $\alpha_n = n!$ and $\beta_n = (n-1)!$. We call this the Poisson check and all our formulae pass it.

We now derive a useful and wider set of checks. By (5·4) it is evident that

$$*\Phi_s(\zeta) = \frac{1}{s}+\frac{\zeta}{s\{1-*F(s)-\zeta\}},$$

so that

$$\sum_{n=0}^{\infty}\frac{\sigma_n\zeta^n}{n!} = 1+\frac{\zeta}{2\pi i}\oint_C\frac{ds}{s\{1-*F(s)-\zeta\}}, \tag{7·3}$$

where C is some circle $|s| = \delta$ in the s-plane.

Let us suppose, that for some positive integer p,

$$*F(s) = 1-s+\frac{(p-1)}{2!}s^2-\frac{(p-1)(p-2)}{3!}s^3+\ldots+\frac{(-s)^p}{p}. \tag{7·4}$$

It is true that there is no distribution function $F(x)$ with the above expression for its L.-S.T., but we are interested only in formal relations between coefficients so this fact need not trouble us. From (7·4) we see that

$$1-*F(s) = \frac{(1-s)^p-1}{-p}$$

and substitution in (7·3) yields

$$\sum_{0}^{\infty}\frac{\sigma_n\zeta^n}{n!} = 1-\frac{\zeta p}{2\pi i}\oint_C\frac{ds}{s\{(1-s)^p-1+p\zeta\}}. \tag{7·5}$$

For ζ arbitrarily small the contour integral (7·5) has two simple poles within C, at $s = 0$ and at $s = 1-(1-p\zeta)^{1/p}$, where the pth root is meant to be the branch taking the value 1 at $\zeta = 0$. A straightforward evaluation of (7·5) by the usual methods of contour integration then gives

$$\sum_{0}^{\infty}\frac{\sigma_n\zeta^n}{n!} = \frac{\zeta}{\{1-(1-p)^{1/p}\}(1-p\zeta)^{1-(1/p)}}.$$

Thus

$$\sum_{1}^{\infty}\frac{\beta_n\zeta^n}{n!} = \log\left\{\sum_{0}^{\infty}\frac{\sigma_n\zeta^n}{n!}\right\} = \log\zeta-\log\{1-(1-p\zeta)^{1/p}\}-\frac{p-1}{p}\log(1-p\zeta).$$

If we differentiate the last equation with respect to ζ we obtain the more manageable result

$$\begin{aligned}\sum_{1}^{\infty}\frac{\beta_n\zeta^{n-1}}{(n-1)!} &= \frac{1}{\zeta}-\frac{\{1-p\zeta\}^{-(p-1)/p}}{1-(1-p\zeta)^{1/p}}+\frac{(p-1)}{1-p\zeta}\\ &= \frac{1}{\zeta}-\frac{1}{p\zeta}[(1-p\zeta)^{-(p-1)/p}+(1-p\zeta)^{-(p-2)/p}+\ldots+(1-p\zeta)^{-1/p}]+\frac{(p-1)}{1-p\zeta}.\end{aligned} \tag{7·6}$$

The right-hand side of (7·6) is in a form suitable for expansion as a power series in ζ, and on comparing coefficients of ζ^{n-1} one deduces that

$$\frac{\beta_n}{(n-1)!} = (p-1)\,p^{n-1} - \frac{1}{n!p}\sum_{r=1}^{p-1} r(r+p)(r+2p)\dots(r+\overline{n-1}p). \tag{7·7}$$

Referring back to (7·4) it appears therefore that our formulae for β_n should give the values indicated by (7·7) when we substitute

$$\left.\begin{aligned} &\mu_1 = 1, \quad \mu_2 = (p-1), \quad \mu_3 = (p-1)(p-2), \\ &\dots \mu_p = (p-1)!, \quad \mu_n = 0 \quad (n > p). \end{aligned}\right\} \tag{7·8}$$

Evidently, by (7·4) we have

$$z_1(\zeta) = 1-(1-p\zeta)^{1/p} = \zeta + \frac{(p-1)}{2!}\zeta^2 + \frac{(p-1)(2p-1)}{3!}\zeta^3 + \dots.$$

If we pick out the coefficient of ζ^n in this expansion we can deduce that

$$\alpha_n = (p-1)(2p-1)\dots(\overline{n-1}p-1). \tag{7·9}$$

Thus the substitution (7·8) should also give (7·9).

We call the checks just described the p-check and all our quoted formulae have passed p-checks for $2 \leqslant p \leqslant 8$. We presume therefore that they are perfectly correct.†

An advantage of the p-check in the computation of the $\{\beta_n\}$ formulae is that the vanishing of the μ_n for $n > p$ enables one to isolate inevitable errors to a few terms. This makes the search for errors easier.

APPENDIX

A combinatorial identity

From Theorem 3 it is possible to deduce a curious combinatorial identity. Let us say that a multipartite number $S \equiv (p_1^{r_1}, p_2^{r_2}, \dots, p_k^{r_k})$ is *allowable* if its degree d is not less than its weight w (one of the $\{p_j\}$ may be zero). If S is allowable then associate with it a coefficient

$$A(S) = \frac{d!}{(d-w)!\,(p_1!)^{r_1}(p_2!)^{r_2}\dots(p_k!)^{r_k}}.$$

A separation $S_1^{q_1}S_2^{q_2}\dots S_m^{q_m}$ of S is *allowable* if every separate S_i is allowable. To an allowable separation associate a coefficient

$$B(S_1^{q_1}S_2^{q_2}\dots S_m^{q_m}) = \frac{(-1)^{q-1}(q-1)!\,[A(S_1)]^{q_1}\dots[A(S_m)]^{q_m}}{q_1!\,q_2!\dots q_m!},$$

where $q = q_1+q_2+\dots+q_m$.

Then the identity in question is that

$$\Sigma^{\#} B(S_1^{q_1}S_2^{q_2}\dots S_m^{q_m}) = 0, \tag{A. 1}$$

whenever the summation $\Sigma^{\#}$ is over all allowable separations of an allowable multipartite number S, of degree d, weight w, with $d > w+1$.

By choosing S in special ways a large number of interesting special cases may be deduced from (A. 1) but we shall not pursue this topic. If we were able to prove (A. 1) by elementary argument we would be in

† I am indebted to Mr D. A. East of the Statistical Laboratory, University of Cambridge, for the calculation of α_8, β_8 and the performance of all these checks. Incidentally, I am also indebted to him for noticing an additional check that no separations of a given multipartite number have been omitted in the working. For the form factors summed over all separations of a given multipartite number total to zero, except when the number is of the form (π_1^i). This can be checked algebraically, for the total involved with a number $(\pi_1^{\rho_1}, \pi_2^{\rho_2}, \dots, \pi_\nu^{\rho_\nu})$ is the coefficient of $x_1^{\rho_1}x_2^{\rho_2}\dots x_\nu^{\rho_\nu}$ in the logarithmic expansion of $\{(1-x_1)(1-x_2)\dots(1-x_\nu)\}^{-1}$.

a position to remove the complex-variable argument needed in the proof of Theorem 3. As it is, we deduce (A. 1) from Theorem 3.

To prove (A.1) suppose that $s^*\phi_1(s) = e_0 + e_1 s + e_2 s^2 + \ldots$.

Then
$$s^d \frac{^*\phi_d(s)}{d!} = \sum_{w=0}^{\infty} A_w^d s^s, \tag{A.2}$$

where
$$A_w^d = d!\, \Sigma^* \frac{e_{\pi_1}^{\rho_1} e_{\pi_2}^{\rho_2} \ldots e_{\pi_\nu}^{\rho_\nu}}{(\pi_1!)^{\rho_1} \ldots (\pi_\nu!)^{\rho_\nu}},$$

and Σ^* has the usual meaning.

If we invert equation (A. 2) we have
$$\frac{\phi_d(t)}{d!} = \sum_{w=0}^{w=d} A_w^d \frac{t^{d-w}}{(d-w)!} + o_d(t), \tag{A.3}$$

where we may assume (by Theorem 2) that the remainder term $o_d(t)$ is negligible. We deduce from (A. 3) that the coefficient of t^r in $\phi_d(t)$ is a linear combination of e-products of degree d and weight $d-r$. It then follows easily that the same remark must also apply to $\psi_d(t)$.

Consider a particular e-product $E \equiv e_0^{\sigma_0} e_1^{\sigma_1} \ldots e_m^{\sigma_m}$, of weight w and degree d, with $\sigma_i \geqslant 0$ (this is to allow some e_i to be absent). Suppose, moreover, that $d > w+1$, so that the associated multipartite number $S \equiv (0^{\sigma_0} 1^{\sigma_1} \ldots m^{\sigma_m})$ is of the type required for (A. 1) to be true. Then the coefficient of t^{d-w} in $\psi_d(t)$ must contain some multiple of E. But since $d-w > 1$, Theorem 3 insists that the coefficient of t^{d-w} must vanish identically. Hence the multiplier of E must be identically zero. This is, essentially, the proof of (A. 1). We must now effect an alternative evaluation of the multiplier, however, to show why this is so.

Let
$$\Pi \equiv S_{11}^{q_{11}} S_{12}^{q_{12}} \ldots S_{1a}^{q_{1a}} S_{21}^{q_{21}} \ldots S_{2b}^{q_{2b}} \ldots S_{k1}^{q_{k1}} \ldots S_{kl}^{q_{kl}}$$

be a separation of S, with S_{ij} of weight w_{ij} and degree d_i for all j. This separation corresponds to a factorization of E into $E_{11}^{q_{11}} E_{12}^{q_{12}} \ldots E_{kl}^{q_{kl}}$, where the e-product E_{ij} is that associated with S_{ij}. We are interested in the way E can appear in $\psi_d(t)$. By (A. 3), assuming $o_d(t)$ negligible, it is clear that E_{ij} cannot appear in $\phi_{d_i}(t)$ unless S_{ij} is allowable. Thus we must suppose Π to be an allowable separation of S, and it is then apparent that the factorization $E_{11}^{q_{11}} \ldots E_{kl}^{q_{kl}}$ must emerge, in $\psi_d(t)$, from the term

$$\frac{(-1)^{q-1}(q-1)!\, \phi_{d_1}^{q_1} \phi_{d_2}^{q_2} \ldots \phi_{d_k}^{q_k}}{q_1!\, q_2! \ldots q_k!\, (d_1!)^{q_1} \ldots (d_k!)^{q_k}}, \tag{A.4}$$

where
$$q_i = \Sigma_j q_{ij} \quad \text{and} \quad q = \Sigma_i q_i.$$

We can obtain $E_{11}^{q_{11}} E_{12}^{q_{12}} \ldots E_{1a}^{q_{1a}}$ from $\phi_{d_1}^{q_1}$ in
$$\frac{q_1!}{q_{11}!\, q_{12}! \ldots q_{1a}!}$$
ways, each time with a coefficient which is, by (A. 3), simply
$$[A(S_{11})]^{q_{11}} [A(S_{12})]^{q_{12}} \ldots [A(S_{1a})]^{q_{1a}}.$$

Similar remarks apply to the other batches of factors of E of equal degree. We can then deduce from (A. 4) that the contribution to the coefficient of E, in $\psi_d(t)/d!$ which corresponds to the separation Π is
$$\frac{(-1)^q (q-1)! [A(S_{11})]^{q_{11}} \ldots [A(S_{kl})]^{q_{kl}}}{q_{11}!\, q_{12}! \ldots q_{kl}!} = B(\Pi).$$

Thus our alternative evaluation of the multiplier of E in $\psi_d(t)/d!$ is simply $\Sigma^\sharp B(\Pi)$, and the identity (A. 1) is proved.

REFERENCES

Adams, E. P. & Hippisley, R. L. (1922). *Smithsonian Mathematical Formulae and Tables of Elliptic Functions, Smithsonian Misc. Coll.* **74** (1). Washington D.C., Smithsonian Institute.

Bartlett, M. S. (1949). Some evolutionary stochastic processes. *J.R. Statist. Soc.* B, **11**, 211–29.

Beneš, V. E. (1957). On queues with Poisson arrivals. *Ann. Math. Statist.* **28**, 670–7.

Blackwell, D. (1948). A renewal theorem. *Duke Math. J.* **15**, 145–50.

Cox, D. R. (1955). Some statistical methods connected with series of events. *J.R. Statist. Soc.* B, **17**, 129–64.

Cox, D. R. & Smith, W. L. (1954). On the superposition of renewal processes. *Biometrika*, **41**, 91–9.

Dwight, H. B. (1957). *Tables of Integrals and other Mathematical Data*, 3rd edition. New York: Macmillan.

FELLER, W. (1941). On the integral equation of renewal theory. *Ann. Math. Statist.* **12**, 243–67.
FELLER, W. (1949). Fluctuation theory of recurrent events. *Trans. Amer. Math. Soc.* **67**, 98–119.
GOURSAT, E. (1933). *Cours d'Analyse Mathématique*, 5th edition. Paris: Gauthier-Villars.
JORDAN, C. (1947). *Calculus of Finite Differences*, 2nd edition. New York: Chelsea.
KENDALL, M. G. & STUART, A. (1958). *The Advanced Theory of Statistics*, **1**. London: Charles Griffin.
PALEY, R. E. A. C. & WIENER, N. (1934). *Fourier Transforms in the Complex Domain, Amer. Math. Soc. Colloquium Publications*, XIX. New York.
SKELLAM, J. G. & SHENTON, L. R. (1957). Distributions associated with random walk and recurrent events. *J.R. Statist. Soc.* B, **19**, 64–118.
SMITH, W. L. (1953). *Transient Queue and Storage Phenomena.* Chapter 7 of Ph.D. Thesis, University of Cambridge.
SMITH, W. L. (1954). Asymptotic renewal theorems. *Proc. Roy. Soc. Edinb.* A, **64**, 9–48.
SMITH, W. L. (1955). Regenerative stochastic processes. *Proc. Roy. Soc.* A, **232**, 6–31.
SMITH, W. L. (1957). On renewal theory, counter problems, and quasi-Poisson processes. *Proc. Camb. Phil. Soc.* **53**, 175–93.
TAKÁCS, L. (1956). On a probability problem arising in the theory of counters. *Proc. Camb. Phil. Soc.* **52**, 488–98.
TÄCKLIND, S. (1945). Fourieranalytische Behandlung vom Erneuerungsproblem. *Skand. AktuarTidskr.* 68–105.
VAN ORSTRAND, C. E. (1910). Reversion of power series. *Phil. Mag.* **19**, 366–76.

A RANDOM WALK IN WHICH THE STEPS OCCUR RANDOMLY IN TIME

By A. MERCER*
Atomic Weapons Research Establishment, Aldermaston
and C. S. SMITH
National Coal Board, London

Summary

The steps of a one-dimensional random walk are positive and occur randomly in time at a fixed mean rate. The sizes of the steps are independent and the size of each step has the same given probability distribution. The distribution of the time to reach a fixed barrier is obtained and approximations to its moments are derived. The results are extended to the case in which the barrier and the random walk process converge at a constant rate.

1. Introduction

The problem considered in this paper arose from the study of the wear of conveyor belting. Incidents which damage belting are of various degrees of severity and different types of incident occur with different frequencies. From a given distribution of the life of belting, information is required about the relative importance of different types of incident. In particular, it is desirable to know whether the wear caused by nearly continuous abrasion is more or less important than the damage due to the more destructive but less frequent blows.

The model, which has been used, assumes that damage accumulates until the total reaches a certain level, when the belt is considered to be worn out completely. The damage done by each incident may be considered as a step of a random walk process in which the steps occur at a constant mean rate, and the problem is to relate the distribution of the size of steps to the distribution of the time taken to reach a barrier.

2. Time to reach a fixed barrier

Consider the one-dimensional random walk process in which the steps are positive and occur randomly at mean rate m, but the size of each step, x, is not specified. Suppose that the sizes of the steps are independent and the probability density of the step size is $f(x)$. Assume that the process starts from the point $u = 0$ at time $t = 0$ and let $q(u, t)$ denote the probability that the distance travelled at time t is less than or equal to u. Then

$$q(u,t) = e^{-mt} \sum_{r=0}^{\infty} \frac{(mt)^r}{r!} p_r(u), \tag{1}$$

where $p_r(u)$ is the probability that the distance covered by r steps is less than or equal to u. Therefore, if $p(t)$ is the probability density of the time to reach a barrier at a distance h from the initial position,

$$p(t) = \frac{d}{dt}\{1 - q(h,t)\} = me^{-mt} \sum_{r=0}^{\infty} \frac{(mt)^r}{r!} \{p_r(h) - p_{r+1}(h)\} \tag{2}$$

* Formerly at the National Coal Board.

and the kth moment of the time about time zero is

$$\frac{k}{m^k}\sum_{r=0}^{\infty}\frac{(r+k-1)!}{r!}p_r(h). \tag{3}$$

But $p_r(u)=\int_0^u p_{r-1}(x)f(u-x)\,dx$ and $p_0(u)=1$ for $u>0$, so that if the Laplace transform with respect to x of $f(x)$ is denoted by $F(s)$ and the Laplace transform with respect to u of $p_r(u)$ is denoted by $P_r(s)$, then

$$P_r(s)=P_{r-1}(s)\,F(s)=\ldots=P_0(s)\,\{F(s)\}^r=\frac{1}{s}\{F(s)\}^r. \tag{4}$$

Therefore the Laplace transform with respect to h of the kth moment of the time about time zero is

$$\frac{k}{m^k}\sum_{r=0}^{\infty}\frac{(r+k-1)!}{r!}\frac{1}{s}\{F(s)\}^r=\frac{k!}{m^k s}\{1-F(s)\}^{-k}. \tag{5}$$

If $a_r=\int_0^{\infty}x^r f(x)\,dx$ is assumed to exist for all r, then

$$F(s)=\int_0^{\infty}e^{-sx}f(x)\,dx=\sum_{r=0}^{\infty}\frac{a_r(-s)^r}{r!}, \tag{6}$$

so that from (5) and (6) as $s\to 0$, the Laplace transforms with respect to h of the first three moments of the time about time zero are

$$\left.\begin{aligned}
&\frac{1}{m}\left\{\frac{1}{a_1 s^2}+\frac{a_2}{2a_1^2 s}+0(1)\right\},\\
&\frac{2}{m^2}\left\{\frac{1}{a_1^2 s^3}+\frac{a_2}{a_1^3 s^2}+\left(\frac{3a_2^2}{4a_1^4}-\frac{a_3}{3a_1^3}\right)\frac{1}{s}+0(1)\right\},\\
&\frac{6}{m^3}\left\{\frac{1}{a_1^3 s^4}+\frac{3a_2}{2a_1^4 s^3}+\left(\frac{3a_2^2}{2a_1^5}-\frac{a_3}{2a_1^4}\right)\frac{1}{s^2}+\left(\frac{a_4}{8a_1^4}-\frac{a_2 a_3}{a_1^5}+\frac{5a_2^3}{4a_1^6}\right)\frac{1}{s}+0(1)\right\}.
\end{aligned}\right\} \tag{7}$$

But it is well known (for example, Widder, 1946) that if $G(s)$ is the Laplace transform of $g(y)$ with respect to y and as $s\to+0$, $G(s)\sim As^{-k}$ where k is non-negative, then under certain conditions, $g(y)\sim A\{\Gamma(k)\}^{-1}y^{k-1}$ as $y\to\infty$. Hence equations (7) suggest that as $h\to\infty$ the mean, variance and third moment about the mean are

$$\left.\begin{aligned}
\mu&=\frac{1}{m}\left\{\frac{h}{a_1}+\frac{a_2}{2a_1^2}+o(1)\right\},\\
\sigma^2&=\frac{1}{m^2}\left\{\frac{a_2 h}{a_1^3}+\frac{5a_2^2}{4a_1^4}-\frac{2a_3}{3a_1^3}+o(1)\right\},\\
\mu_3&=\frac{1}{m^3}\left\{\left(\frac{3a_2^2}{a_1^5}-\frac{a_3}{a_1^4}\right)h+\frac{11a_2^3}{2a_1^6}-\frac{5a_2 a_3}{a_1^5}+\frac{3a_4}{4a_1^4}+o(1)\right\}.
\end{aligned}\right\} \tag{8}$$

Alternatively, the $(i+1)$th step may be considered as the ith renewal of a process and then a rigorous proof of the asymptotic expressions for the moments can be obtained using results in renewal theory (Smith, 1959).

3. EXAMPLES

The following three examples illustrate the preceding results.

(i) Suppose that the size of step has a negative exponential distribution, $f(x)=\beta e^{-\beta x}$. Then

$$F(s)=\int_0^{\infty}\beta e^{-(s+\beta)x}\,dx=\frac{\beta}{\beta+s}.$$

Also $p_r(u)$ can be obtained directly as $e^{-\beta u}\sum_{q=r}^{\infty}\frac{(\beta u)^q}{q!}$ and therefore from (2),

$$p(t) = me^{-(m\lambda+\beta h)}\sum_{r=0}^{\infty}\frac{(mt)^r}{r!}\frac{(\beta h)^r}{r!}. \tag{9}$$

Hence the mean, variance and third moment about the mean of the times taken to reach the barrier are

$$\left.\begin{aligned}\mu &= \frac{1}{m}(1+\beta h),\\ \sigma^2 &= \frac{1}{m^2}(1+2\beta h),\\ \mu_3 &= \frac{2}{m^3}(1+3\beta h).\end{aligned}\right\} \tag{10}$$

But $a_r = \int_0^\infty \beta x^r e^{-\beta x}\,dx = \frac{r!}{\beta^r}$ and substituting in (8) we see that the asymptotic results are exact.

(ii) Suppose that the size of step is the sum of two independent variables which are identically distributed according to the negative exponential distribution.

Then

$$f(x) = \beta^2 x e^{-\beta x} \quad \text{and} \quad F(s) = \left(\frac{\beta}{\beta+s}\right)^2.$$

Also $p_r(u)$ can be obtained directly as

$$e^{-\beta u}\sum_{q=2r}^{\infty}\frac{(\beta u)^q}{q!}$$

and therefore from (2)

$$p(t) = me^{-(m\lambda+\beta h)}\sum_{r=0}^{\infty}\frac{(mt)^r}{r!}\left\{\frac{(\beta h)^{2r}}{(2r)!}+\frac{(\beta h)^{2r+1}}{(2r+1)!}\right\} \tag{11}$$

so that as $h \to \infty$,

$$\left.\begin{aligned}\mu &= \frac{1}{m}\left\{\frac{3}{4}+\frac{\beta h}{2}+0(e^{-2\beta h})\right\},\\ \sigma^2 &= \frac{1}{m^2}\left\{\frac{13}{16}+\frac{3\beta h}{4}+0(\beta h\, e^{-2\beta h})\right\},\\ \mu_3 &= \frac{3}{m^3}\left\{\frac{9}{16}+\frac{5\beta h}{8}+0(\beta^2 h^2 e^{-2\beta h})\right\}.\end{aligned}\right\} \tag{12}$$

But $a_r = (r+1)!/\beta^r$ and on substituting in (8) we see that the asymptotic results only differ from the exact results by terms containing negative exponentials.

(iii) Suppose all the steps are of equal size ξ. Then $2mt$ has the χ^2-distribution with $2n$ degrees of freedom where n is an integer satisfying $n \geqslant h/\xi > n-1$ and $\mu = n/m$, $\sigma^2 = n/m^2$ and $\mu_3 = 2n/m^3$. On substituting $a_r = \xi^r$ in (8), we see that the asymptotic values of these three moments are

$$\left.\begin{aligned}\mu &= \frac{1}{m}\left(\frac{h}{\xi}+\frac{1}{2}\right),\\ \sigma^2 &= \frac{1}{m^2}\left(\frac{h}{\xi}+\frac{7}{12}\right),\\ \mu_3 &= \frac{2}{m^3}\left(\frac{h}{\xi}+\frac{5}{8}\right).\end{aligned}\right\} \tag{13}$$

4. TIME TO REACH A MOVING BARRIER

Now suppose that the barrier is moving, relative to the random walk process, at a fixed rate λ per unit time towards the origin. Then at time t, the barrier is at a distance $(h-\lambda t)$ from the origin and the process is certain to have terminated by time h/λ. In the application to the life of conveyor belting, the relative movement is due to continuous wear of the belting, which occurs in addition to the random blows. This wear is the limiting case of that in which very small blows occur frequently. Hence, suppose that the steps consist of those whose size has probability density $f(x)$ and occur at mean rate m and steps of size x_1 occurring at mean rate m_1. Then if $f_1(x)$ is the density function of the step size distribution

$$(m+m_1)f_1(x) = mf(x)+m_1\delta(x-x_1).* \tag{14}$$

Therefore
$$(m+m_1)\gamma_r = ma_r+m_1x_1^r, \tag{15}$$

where
$$\gamma_r = \int_0^\infty x^r f_1(x)\,dx$$

and
$$a_r = \int_0^\infty x^r f(x)\,dx$$

as previously. Now let $m_1 \to \infty$ and $x_1 \to 0$ so that $m_1x_1 \to \lambda$. Then from (15),

$$(m+m_1)\gamma_1 \to ma_1\left(1+\frac{1}{\rho}\right) \tag{16}$$

and
$$(m+m_1)\gamma_r \to ma_r \quad \text{for} \quad r \geqslant 2, \tag{17}$$

where
$$\rho = \frac{ma_1}{\lambda}. \tag{18}$$

Now the asymptotic values of the mean, variance and third moment about the mean of the distribution of the time to reach the barrier when the step-size distribution has density function $f_1(x)$ are given by replacing a_r by γ_r in (8). Although it has not been proved rigorously, these results suggest that if (16), (17) and (18) are then used, the asymptotic values of the mean, variance and third moment about the mean when the barrier and random walk process converge at a constant rate are,

$$\left.\begin{aligned}
\mu &= \frac{1}{m}\left(\frac{\rho}{1+\rho}\right)^2\left\{\frac{h}{a_1}\left(\frac{1+\rho}{\rho}\right)+\frac{a_2}{2a_1^2}\right\},\\
\sigma^2 &= \frac{1}{m^2}\left(\frac{\rho}{1+\rho}\right)^4\left\{\frac{a_2h}{a_1^3}\left(\frac{1+\rho}{\rho}\right)+\frac{5a_2^2}{4a_1^4}-\frac{2a_3}{3a_1^3}\left(\frac{1+\rho}{\rho}\right)\right\},\\
\mu_3 &= \frac{1}{m^3}\left(\frac{\rho}{1+\rho}\right)^6\left\{\left[\frac{3a_2^2}{a_1^5}-\frac{a_3}{a_1^4}\left(\frac{1+\rho}{\rho}\right)\right]\left(\frac{1+\rho}{\rho}\right)h+\frac{11a_2^3}{2a_1^6}-\frac{5a_2a_3}{a_1^5}\left(\frac{1+\rho}{\rho}\right)+\frac{3a_4}{4a_1^4}\left(\frac{1+\rho}{\rho}\right)^2\right\}.
\end{aligned}\right\} \tag{19}$$

5. PRACTICAL APPLICATION

In the previous sections the distribution of the time to reach a barrier has been found. Expressions for the moments of that time have been obtained as functions of the mean frequency of the steps, the moments of the step-size distribution, the distance of the barrier from the origin and the rate at which the barrier and random walk process converge. In practice, the moments of the life distribution are known and the problem is to make some inference about the discrete blows and continuous wear.

[* For definition of $\delta(h)$ see footnote to p. 68. ED.].

If the variation in the size of the blows is negligible, then from (19) the asymptotic expressions for the first three moments are

$$\left.\begin{aligned}\mu &= \frac{1}{m}\left(\frac{\rho}{1+\rho}\right)^2\left\{\left(\frac{1+\rho}{\rho}\right)\frac{h}{\xi}+\frac{1}{2}\right\},\\ \sigma^2 &= \frac{1}{m^2}\left(\frac{\rho}{1+\rho}\right)^4\left\{\left(\frac{1+\rho}{\rho}\right)\frac{h}{\xi}+\frac{5}{4}-\frac{2}{3}\left(\frac{1+\rho}{\rho}\right)\right\},\\ \mu_3 &= \frac{1}{m^3}\left(\frac{\rho}{1+\rho}\right)^6\left\{\left(2-\frac{1}{\rho}\right)\left(\frac{1+\rho}{\rho}\right)\frac{h}{\xi}+\frac{11}{2}-5\left(\frac{1+\rho}{\rho}\right)+\frac{3}{4}\left(\frac{1+\rho}{\rho}\right)^2\right\},\end{aligned}\right\} \tag{20}$$

where ξ is the size of the discrete blows. Hence $\rho = m\xi/\lambda$ is the relative importance of the discrete blows to that of the continuous wear. This is the parameter in which one is most interested in order to decide whether the surface of the conveyor belting, which is largely affected by the continuous abrasion, or the carcase, which has to stand heavy infrequent blows, is in greater need of improvement. The exact expressions for the moments are too

Table 1. *Exact and asymptotic values of moments*

ρ	$m\mu$		$m^2\sigma^2$		$m^3\mu_3$	
	Exact	Asymptotic	Exact	Asymptotic	Exact	Asymptotic
0·5	1·056	1·056	0·102	0·102	−0·004	−0·004
1·0	1·625	1·625	0·370	0·370	0·075	0·070
2·0	2·229	2·222	0·901	0·938	0·620	0·565
10·0	2·997	3·140	2·956	2·607	5·455	4·052
∞	3·000	3·500	3·000	3·589	6·000	7·250

complicated to be used for estimating the parameters of the distribution, but it can be shown numerically that convergence is rapid if ρ is not large. If h/ξ is an integer, n say, then it can be shown directly that the first three moments about the origin are exactly

$$\left.\begin{aligned}\mu &= \frac{1}{m}\left\{n-\sum_{q=1}^{n}\sum_{r=0}^{n-q}\frac{(\rho q)^r}{r!}e^{-\rho q}\right\},\\ \mu_2' &= \frac{1}{m^2}\left\{n(n+1)-2\sum_{q=1}^{n}\sum_{r=0}^{n-q+1}(n-q+1)\frac{(\rho q)^r}{r!}e^{-\rho q}\right\},\\ \mu_3' &= \frac{1}{m^3}\left\{n(n+1)(n+2)-3\sum_{q=1}^{n}\sum_{r=0}^{n-q+2}(n-q+1)(n-q+2)\frac{(\rho q)^r}{r!}e^{-\rho q}\right\}.\end{aligned}\right\} \tag{21}$$

The exact and asymptotic values of the mean, variance and third moment about the mean are given in Table 1 for $h/\xi = 3$. It will be seen that the asymptotic value for the variance oscillates about the exact value. Now let $\rho = (x-1)^{-1}$, $\mu m = x^{-2}y$ and $h/\xi = x^{-1}z$ and denote σ^2/μ^2 by α_1 and μ_3/μ^3 by α_2. Then the three equations given by the asymptotic expressions for the moments are

$$\left.\begin{aligned}y &= \tfrac{1}{2}+z,\\ \alpha_1 y^2 &= \tfrac{5}{4}-\tfrac{2}{3}x+z,\\ \alpha_2 y^3 &= \tfrac{11}{2}-5x+\tfrac{3}{4}x^2+(3-x)z.\end{aligned}\right\} \tag{22}$$

Hence estimates of ρ, m and h/ξ follow from the solution of

$$432\alpha_1^2 y^4 - 32(15\alpha_1 + 8\alpha_2)y^3 + 24(2 + 45\alpha_1)y^2 - 600y - 29 = 0. \quad (23)$$

In the course of our investigation equations (22) and (23) were used to calculate a table of values of ρ, μm and h/ξ for certain combinations of values of α_1 and α_2. We should be glad to make these data available to any reader of this paper who wishes to go further into the application of the theory.

It would be possible to estimate the standard errors of the estimates of x, y and z and hence those of ρ, m and h/ξ. However, the formulae are too complicated for any assessment of the efficiency to be practical but, in any case, the only possible method of estimation appears to be that of moments.

This paper is published by permission of the National Coal Board.

REFERENCES

SMITH, W. L. (1959). On the cumulants of renewal processes. *Biometrika*, **46**, 1–29.
WIDDER, D. V. (1946). *The Laplace Transform*. Princeton University Press.

A TEST OF HOMOGENEITY FOR ORDERED ALTERNATIVES

By D. J. BARTHOLOMEW
University College of North Staffordshire

1. Introduction

A problem which occurs frequently in statistical analysis is that of testing whether a number of normal variates have the same mean. Given k such variates with mean values $m_1, m_2, \ldots, m_k$ the null hypothesis which is tested may be written,

$$H_0: m_1 = m_2 = \ldots = m_k.$$

The object of this paper is to give a new solution to this problem which has many applications, especially in the social sciences. The appropriate test of the hypothesis can only be determined when the alternative hypotheses, which are considered likely or important, have been specified. The most general form of alternative, which is the basis of almost all existing work, may be written,

$$H_1: m_1 \neq m_2 \neq \ldots \neq m_k.$$

As examples we note that the standard analysis of variance tests fall into this category. The χ^2-test for the equality of proportions in a $2 \times k$ contingency table is a further example if the numbers involved are large enough. In some applications it is possible to be more precise in the specification of the alternative. When this is the case it is advantageous to make use of this extra information to obtain more powerful tests. This paper deals with what have been called 'ordered' alternatives for which the rank order of the means is known. Thus instead of H_1 above, we shall consider

$$H_2: m_1 \geqslant m_2 \geqslant \ldots \geqslant m_k.$$

At least one of the inequality signs must be a strict inequality so that H_0 is not included in H_2. The practical importance of this alternative will be clear from the discussion below. For the simple case $k = 2$ the situation is met by using a single-tail test, but for $k > 2$ the distinction between one- and two-tail tests is lost. The present work may therefore be looked upon as a generalization of the single-tail test.

Attempts to meet the need for a test against ordered alternatives have already been made. Jonckheere (1954) discussed the one-way analysis of variance and gave a distribution-free test which is closely related to Kendall's coefficient of rank correlation. The need for a test of this kind arises in psychological work where qualitative characteristics can be ranked but not easily measured. Jonckheere mentions, as an example, an experiment to test the effect of stress on a task of manual dexterity. Here data would be obtained from groups of subjects working under high, medium, low and minimal stress, the null hypothesis being that stress has no effect on performance, and the alternative that increasing stress produces an increasing effect. Many similar examples occur in the analysis of social survey data where people are often grouped according to 'social class', 'living conditions', etc. Such factors can seldom be measured satisfactorily but they can be ranked according to their expected effect.

Armitage (1955), who discussed the problem in connexion with the analysis of $2 \times k$ contingency tables, drew his examples from the medical field. He was particularly interested

in linear changes in the proportions and gave tests which amount to converting the problem to one of regression, by allocating scores to the experimental groups.

Neither of the approaches to the problem mentioned above provides a general solution, since each was designed for a particular application. The new tests to be derived are based on the likelihood ratio and are applicable to any problem which can be formulated as a test for the equality of normally distributed means. They are closely related to the familiar χ^2 and F criteria which are used when the alternative is H_1. Significance is easily tested using tabulated functions or the special tables of percentage points included in the paper.

2. THE TEST CRITERIA

2·1. Suppose that we have independent observations $x_1, x_2, \ldots, x_k$, where x_i is normally distributed with mean m_i and standard deviation σ_i $(i = 1, 2, \ldots, k)$. It is important to be clear on the meaning of the suffices; thus x_1 is the observation from the population with the largest mean under H_2 but is not necessarily the largest observation. In this section it will be assumed that the σ's are known, but in § 2·2 we shall include the case when $\sigma_i = \sigma/\sqrt{n_i}$ $(i = 1, 2, \ldots, k)$, the n's being known sample numbers and σ being estimated from the data.

Our problem is to construct a test criterion which is particularly sensitive to the alternative H_2. Since the χ^2-test, used when the alternative is H_1, can be derived by the likelihood ratio principle it seems reasonable to use the same approach when it is H_2. Although it is not clear, in this case, whether the test derived has optimum properties, the method has a strong intuitive appeal and leads to a meaningful test.

The standard method of deriving likelihood ratio tests leads to the following form of test criterion

$$\overline{\chi}_k^2 = \sum_1^k a_i(x_i - \bar{x})^2 - \sum_1^k a_i(x_i - \hat{m}_i)^2, \tag{1}$$

where $a_i = 1/\sigma_i^2$, $\bar{x} = \sum_1^k a_i x_i \Big/ \sum_1^k a_i$ and $\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_k$ are the values of $m_1, m_2, \ldots, m_k$ which maximize (1) subject to the condition $\hat{m}_1 \geqslant \hat{m}_2 \geqslant \ldots \geqslant \hat{m}_k$. The symbol $\overline{\chi}_k^2$ is used because of the close link with the usual χ^2-test which will become apparent in what follows. The effect of restricting the alternative to H_2 is to introduce the second term on the right. The values of $m_1, m_2, \ldots, m_k$ which maximize (1) are those which minimize

$$U_k = \sum_1^k a_i(x_i - m_i)^2,$$

so the problem of determining $\overline{\chi}_k^2$ has been reduced to the following. Minimize U_k with respect to $m_1, m_2, \ldots, m_k$, subject to the restriction $m_1 \geqslant m_2 \geqslant \ldots \geqslant m_k$.

It will simplify the determination of the minimum of U_k if we use geometrical terminology. The first step is to discover whether the absolute minimum of U_k lies in the region H_2 $(m_1 \geqslant m_2 \geqslant \ldots \geqslant m_k)$. If this is so the restricted and unrestricted problems have the same solution, which, in this case, is obviously $\hat{m}_i = x_i$ for $i = 1, 2, \ldots, k$. This solution lies in H_2 if $x_1 \geqslant x_2 \geqslant \ldots, \geqslant x_k$.

If this is not the case then it will be possible to find $x_j < x_{j+1}$ for some j. The absolute minimum would require $\hat{m}_j < \hat{m}_{j+1}$ which is outside H_2. To avoid this we must introduce the restriction $m_j \geqslant m_{j+1}$ before minimizing U_k, as follows.

$$\begin{aligned} U_k &= \sum_{i \neq j, j+1} a_i(x_i - m_i)^2 + \{a_j(x_j - m_j)^2 + a_{j+1}(x_{j+1} - m_{j+1})^2\} \\ &= R + S, \quad \text{say.} \end{aligned}$$

The minimum of R is unaffected by the restriction and is therefore zero when $\hat{m}_i = x_i$. Consider now the minimum of S. The point $m_j = x_j$, $m_{j+1} = x_{j+1}$ falls outside the region $m_j \geqslant m_{j+1}$ in the (m_j, m_{j+1}) plane. The minimum of S will therefore occur on the boundary of the region, that is when $m_j = m_{j+1}$. In this case it is easily shown that

$$\hat{m}_j = \hat{m}_{j+1} = \frac{a_j x_j + a_{j+1} x_{j+1}}{a_j + a_{j+1}}.$$

It will be convenient, for what follows, to introduce the notation

$$X(p,q) = \sum_{i=p}^{q} a_i x_i \Big/ \sum_{i=p}^{q} a_i.$$

The minimum of U_k subject to the restriction $m_j \geqslant m_{j+1}$ may therefore be written

$$\hat{m}_j = \hat{m}_{j+1} = X(j, j+1),$$
$$\hat{m}_i = x_i \quad (i \neq j, j+1).$$

The problem has now been reduced to one involving $(k-1)$ variates. If

$$x_1 \geqslant x_2 \geqslant \ldots \geqslant x_{j-1} \geqslant X(j,j+1) \geqslant x_{j+2} \geqslant \ldots \geqslant x_k,$$

the final solution has been reached. Otherwise the problem is the same as the original but with the number of variates reduced by one. It may therefore be treated in the same way by taking $X(j,j+1)$ as a single observation with weight $(a_j + a_{j+1})$. The rule for obtaining $\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_k$ may, therefore, be stated as follows.

Arrange the observations in the order predicted by the alternative, thus, $x_1, x_2, \ldots, x_k$. If any consecutive pair x_j, x_{j+1} are not in the expected order (i.e. if $x_{j+1} > x_j$), form the average

$$X(j,j+1) = (a_j x_j + a_{j+1} x_{j+1})/(a_j + a_{j+1}).$$

The series is now reduced to $(k-1)$ members of which $(k-2)$ are unchanged and one is the average of two original ones. Proceed exactly as before, treating $X(j,j+1)$ as a single observation but with weight $(a_j + a_{j+1})$. Continue in this manner until the resulting values are in the order predicted by H_2. The following example illustrates the method.

Example

x_1	x_2	x_3	x_4
60	-73	-141	7

$$a_1 = a_2 = a_3 = a_4.$$

We notice that $x_4 > x_3$ so we form their average,

$$X(3,4) = \tfrac{1}{2}(-141+7) = -67.$$

Now $X(3,4) > x_2$ so we calculate

$$\tfrac{1}{3}\{2X(3,4) + x_2\} = X(2,4) = \tfrac{1}{3}(-73-141+7) = -69.$$

As $x_1 > X(2,4)$ the solution is

$$\hat{m}_1 = 60, \quad \hat{m}_2 = \hat{m}_3 = \hat{m}_4 = -69.$$

By calculating $\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_k$ in the way described $\overline{\chi}^2_k$ can be obtained by substitution in (1). It is easily shown, for estimates of the form given by this method that

$$\overline{\chi}^2_k = \sum_1^k a_i (x_i - \bar{x})^2 - \sum_1^k a_i (x_i - \hat{m}_i)^2 = \sum_1^k a_i (\hat{m}_i - \bar{x})^2. \tag{2}$$

The difference between the method of calculation of the usual χ^2 and $\overline{\chi}^2_k$ is that for the latter we use the estimates $\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_k$ instead of the original observations $x_1, x_2, \ldots, x_k$. This is illustrated by the following example.

Example

i	1	2	3	4
x_i	−0·83	1·88	0·20	0·44
a_i	5	10	10	5

Carrying out the procedure described above we find

$$\hat{m}_1 = \hat{m}_2 = 0{\cdot}98, \quad \hat{m}_3 = \hat{m}_4 = 0{\cdot}28, \quad \bar{x} = 0{\cdot}63.$$

Hence $$\overline{\chi}^2_4 = 5(0{\cdot}98)^2 + 10(0{\cdot}98)^2 + 10(0{\cdot}28)^2 + 5(0{\cdot}28)^2 - 30(0{\cdot}63)^2 = 3{\cdot}68.$$

The example also indicates an alternative method of calculation. We could look upon the problem as involving two observations, $X(1,2) = 0{\cdot}98$, $X(3,4) = 0{\cdot}28$ each with weight $(5+10)$. $\overline{\chi}^2_4$ would then be calculated as the usual χ^2 sum of squares using $X(1,2)$, $X(3,4)$ and the appropriate weights. This alternative method is always available and is, in fact, the simplest method of calculation. When the averaging process is complete we have l means,

$$X(1,k_1), \quad X(k_1+1, k_1+k_2), \ldots, X(k-k_l+1, k)$$

with weights $$A(1,k_1), \quad A(k_1+1, k_1+k_2), \ldots, A(k-k_l+1, k),$$

where $$A(p,q) = \sum_p^q a_i, \quad \sum_1^l k_i = k.$$

$\overline{\chi}^2_k$ is then obtained as the sum of squares of these means. This will be referred to as the reduced form of the problem.

2·2. In this section we give the modifications necessary when σ can be estimated from the data. Consider the problem as it arises in the one-way analysis of variance by groups where we have observations $x_{ij}\,(i = 1, 2, \ldots, k;\ j = 1, 2, \ldots, n_i)$, which are normally and independently distributed about a mean m_i with common standard deviation σ. Define

$$\bar{x}_{i.} = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij} \quad \text{and} \quad \bar{x}_{..} = \sum_1^k n_i \bar{x}_{i.} \Big/ \sum_1^k n_i,$$

then we have to test whether $\bar{x}_{1.}, \bar{x}_{2.}, \ldots \bar{x}_{k.}$ come from populations with the same mean value. The mean $\bar{x}_{i.}$ has standard deviation $\sigma/\sqrt{n_i}$ and therefore $a_i = n_i/\sigma^2$. The most obvious modification of the test criterion is the replacement of σ^2 by the residual mean square obtained from the analysis of variance t ble. If this is denoted by S_e^2 the criterion would become

$$\sum_1^k n_i(\bar{x}_{i.} - \bar{x}_{..})^2 / S_e^2.$$

However, a slight improvement can be obtained as follows. Suppose, for example, that $\bar{x}_{i.} < \bar{x}_{i+1.}$, then following the rule for calculating the test criterion we form

$$X(i, i+1) = (n_i \bar{x}_{i.} + n_{i+1}\bar{x}_{i+1.})/(n_i + n_{i+1}).$$

This is exactly equivalent to pooling the ith and $(i+1)$th groups and then calculating the average of the single group so formed. The residual mean square calculated from this new table will have one more degree of freedom than that from the original table, because the 'between groups' degrees of freedom have been decreased by one with the total remaining

constant. In general, it can be shown that if the reduced form of the problem consists of l groups then the increase in the residual degrees of freedom will be $(k-l)$. The test criterion to be used is therefore defined as

$$\overline{F}_k = \sum_1^k n_i(\hat{m}_i - \bar{x}_{..})^2/(l-1)\,S_e'^2, \tag{3}$$

where $S_e'^2$ is the residual mean square of the analysis of variance table obtained from the reduced form of the problem, having $(N-l)$ degrees of freedom $\left(N = \sum_1^k n_i\right)$. $\overline{F}_k$ is thus the ordinary mean square ratio calculated for the reduced problem. An example, illustrating the method of calculation, is included in § 5. It can readily be shown that $\overline{F}_k$ is the criterion produced by the likelihood ratio principle.

3. The distribution of $\overline{\chi}_k^2$ and $\overline{F}_k$

3·1. *The distribution of* $\overline{\chi}_k^2$. To illustrate the method of approach to the distribution problem we consider first the case $k = 2$. This must, of course, reduce to the familiar one-tail test. We have

$$\overline{\chi}_2^2 = \sum_1^2 a_i(x_i - \bar{x})^2 \quad \text{for} \quad x_1 > x_2$$

$$= 0 \qquad\qquad\quad \text{for} \quad x_2 \geqslant x_1.$$

For tests of significance the probability required is

$$\Pr\{\overline{\chi}_2^2 \geqslant \gamma\} = \Pr\left\{x_1 > x_2;\ \sum_1^2 a_i(x_i - \bar{x})^2 \geqslant \gamma\right\}$$

when the null hypothesis is true. Now the value of $\sum_1^2 a_i(x_i-\bar{x})^2$ does not depend on the order of x_1 and x_2 but only on their absolute difference so we may write

$$\Pr\{\overline{\chi}_2^2 \geqslant \gamma\} = \Pr\{x_1 > x_2\}\Pr\left\{\sum_1^2 a_i(x_i - \bar{x})^2 \geqslant \gamma\right\}.$$

Further, since on the null hypothesis $\Pr\{x_1 > x_2\} = \frac{1}{2}$ whatever a_1, a_2, we have finally

$$\Pr\{\overline{\chi}_2^2 \geqslant \gamma\} = \tfrac{1}{2}\Pr\{\chi_1^2 \geqslant \gamma\},$$

where χ_ν^2 denotes χ^2 with ν degrees of freedom. This will be recognized as the equivalent of the one-tail test.

For larger k the method is similar. Let $P(l, k; a_1, a_2, \ldots, a_k)$ denote the probability that the reduced form of the problem consists of l means when the original observations $x_1, x_2, \ldots, x_k$ had weights $a_1, a_2, \ldots, a_k$. When there is no risk of confusion we shall suppress the a's and write $P(l, k)$.

The probability integral of $\overline{\chi}_k^2$ may thus be written

$$\left.\begin{aligned} \Pr\{\overline{\chi}_k^2 \geqslant \gamma\} &= \sum_{l=2}^k P(l,k)\Pr\{\overline{\chi}_k^2 \geqslant \gamma \mid l\}, \\ \Pr\{\overline{\chi}_k^2 = 0\} &= P(1,k), \end{aligned}\right\} \tag{4}$$

where $\Pr\{\overline{\chi}_k^2 \geqslant \gamma \mid l\}$ is the probability that $\overline{\chi}_k^2$ exceeds γ, given that it was calculated from means. This is simply equal to $\Pr\{\chi_{l-1}^2 \geqslant \gamma\}$ as we now show.

An obvious simplification of the notation introduced at the end of § 2·1 gives

$$\overline{\chi}_k^2 = \sum_1^l A_i(X_i - \overline{X})^2,$$

where $$\overline{X} = \sum_1^l A_i X_i / \sum_1^l A_i \quad \text{and} \quad \text{var}(X_i) = A_i^{-1} \quad (i = 1, 2, \ldots, l).$$

Two conditions must be satisfied if $\overline{\chi}_k^2$, for given l, is to be distributed like χ^2 with $(l-1)$ degrees of freedom.

(i) Each X_i $(i = 1, 2, \ldots, l)$ must be normally distributed about a common mean.

(ii) The restriction $X_1 \geqslant X_2 \geqslant \ldots \geqslant X_l$ must be irrelevant as far as the distribution of $\overline{\chi}_k^2$ is concerned. This amounts to showing that the X's may be treated as if they were independent.

The first condition is satisfied, on the null hypothesis, since the averaging which leads to any X_i does not depend on the magnitude of the values which are averaged. They can therefore be regarded as being selected at random.

To show that the second condition is also satisfied, consider the following, algebraically equivalent, expression for $\overline{\chi}_k^2$.

$$\overline{\chi}_k^2 = \sum_{i>j}\sum A_i A_j (X_i - X_j)^2 \Big/ \sum_{i=1}^l A_i.$$

If the X's were independent then $(X_i - X_j)^2$ would be the square of a normal variate with zero mean and $\overline{\chi}_k^2$ would be a linear function of such variates. When the X's are ordered, $(X_i - X_j)$ is distributed like the modulus of a normal variable (since $X_i \geqslant X_j$ when $i > j$). However, the square of the modulus has exactly the same distribution as the square of the difference when the sign is not restricted. The distribution of $\overline{\chi}_k^2$ is therefore unaffected by the dependence among the X's resulting from the restriction $X_1 \geqslant X_2 \geqslant \ldots \geqslant X_l$.

(4) may therefore be re-written

$$\left.\begin{aligned} &\Pr\{\overline{\chi}^2 \geqslant \gamma\} = \sum_2^k P(l,k) \int_\gamma^\infty p(\chi_{l-1}^2)\, d\chi_{l-1}^2, \\ &\Pr(\overline{\chi}^2 = 0\} = P(1,k). \end{aligned}\right\} \tag{5}$$

The distribution problem thus centres on the determination of the probabilities $P(l,k)$. In what follows we give complete results for $k = 3$ and $k = 4$, and a partial result for $k = 5$, using known results about the normal multivariate integral. The following notation is used

$$z_i = x_i - x_{i+1} \quad (i = 1, 2, \ldots . (k-1)).$$

ρ_{ij} is the product moment correlation coefficient of z_i and z_j, which is found to be

$$\rho_{i,i+1} = -\sqrt{\left\{\frac{a_i a_{i+2}}{(a_i + a_{i+1})(a_{i+1} + a_{i+2})}\right\}} \quad \text{for} \quad i = 1, 2, \ldots, (k-1),$$

$$\left.\begin{aligned} \rho_{ij} &= 1 \quad \text{for} \quad i = j, \\ &= 0 \quad \text{for} \quad |i-j| > 1. \end{aligned}\right\} \tag{6}$$

$k = 3$ $$P(3,3) = \Pr\{z_1 > 0, z_2 > 0\}.$$

On the null hypothesis z_1 and z_2 have a normal bivariate distribution with zero means and correlation coefficient ρ_{12}; hence using the well-known result of Sheppard

$$P(3,3) = \frac{1}{2}\left[1 - \frac{\cos^{-1}\rho_{12}}{\pi}\right].$$

In the special case $a_1 = a_2 = a_3, \rho_{12} = -\frac{1}{2}$ and $P(3,3) = 1/3!$; a result which could have been obtained directly since, in this case, all permutations of the x's are equally likely. $P(2,3)$ can be obtained using the known results for $k = 2$. Two possibilities must be considered:

$$\text{(i)} \quad x_1 < x_2, \quad X(1,2) > x_3$$
$$\text{(ii)} \quad x_2 < x_3, \quad X(2,3) < x_1.$$

Now $X(1,2)$ is normally distributed with zero mean and variance $(a_1+a_2)^{-1}$, so that the event $x_1 < x_2$ is independent of the event $X(1,2) > x_3$. Hence the first of these possibilities has probability

$$P(1,2;a_1,a_2)\,P(2,2;(a_1+a_2),a_3).$$

By a similar argument the probability of the second is

$$P(1,2;a_2,a_3)\,P(2,2;a_1,(a_2+a_3)).$$

Because $P(1,2) = P(2,2) = \frac{1}{2}$ whatever the weights,

$$P(2,3) = \tfrac{1}{4}+\tfrac{1}{4} = \tfrac{1}{2}.$$

Using the fact that $\sum_1^3 P(l,k) = 1$, $P(1,3)$ is found to be

$$\frac{1}{2\pi}\cos^{-1}\rho_{12}.$$

$k = 4$ $\qquad P(4,4) = \Pr\{z_1 > 0, z_2 > 0, z_3 > 0\}.$

When z_1, z_2, z_3 are distributed in the trivariate normal form with zero means, David (1953) gives the result

$$P(4,4) = \frac{1}{4\pi}(2\pi - \cos^{-1}\rho_{12} - \cos^{-1}\rho_{23} - \cos^{-1}\rho_{13})$$

the ρ_{ij}'s being given by (6).

$P(3,4)$ is built up, by a similar argument to that used for $k = 3$, using the results for $k = 2$ and 3 as follows:

$$P(3,4) = P(1,2;a_1,a_2)\,P(3,3;(a_1+a_2),a_3,a_4) + P(1,2;a_2,a_3)\,P(3,3;a_1,(a_2+a_3),a_4) + P(1,2;a_3,a_4)\,P(3,3;a_1,a_2,(a_3+a_4)).$$

Therefore $$P(3,4) = \frac{1}{4}\left[\left(1-\frac{\cos^{-1}\rho'_{12}}{\pi}\right)+\left(1-\frac{\cos^{-1}\rho''_{12}}{\pi}\right)+\left(1-\frac{\cos^{-1}\rho'''_{12}}{\pi}\right)\right].$$

The correlations, which are given below, prove to be equal to the partial correlations between the z's, as indicated.

$$\rho'_{12} = -\sqrt{\frac{(a_1+a_2)\,a_4}{(a_1+a_2+a_3)\,(a_3+a_4)}} = \rho_{23\cdot1}$$

$$\rho''_{12} = -\sqrt{\frac{a_1 a_4}{(a_1+a_2+a_3)\,(a_2+a_3+a_4)}} = \rho_{13\cdot2}$$

$$\rho'''_{12} = -\sqrt{\frac{a_1(a_3+a_4)}{(a_1+a_2)\,(a_2+a_3+a_4)}} = \rho_{12\cdot3},$$

whence $$P(3,4) = \tfrac{3}{4} - (\cos^{-1}\rho_{12\cdot3} + \cos^{-1}\rho_{23\cdot1} + \cos^{-1}\rho_{13\cdot2})/4\pi.$$

In the same manner the equation for $P(2,4)$ is found to be

$$P(2,4) = P(1,3;a_1,a_2,a_3)\,P(2,2;(a_1+a_2+a_3),a_4) + P(1,3;a_2,a_3,a_4)\,P(2,2;a_1,(a_2+a_3+a_4)) + P(1,2;a_1,a_2)\,P(1,2;a_3,a_4)\,P(2,2;(a_1+a_2),(a_3+a_4)).$$

Substituting known results in this equation we find

$$P(2,4) = \tfrac{1}{8} + [\cos^{-1}\rho_{12} + \cos^{-1}\rho_{23}]/4\pi$$
$$= \tfrac{1}{2} - P(4,4).$$

Finally, by subtraction

$$P(1,4) = (\cos^{-1}\rho_{12\cdot3} + \cos^{-1}\rho_{23\cdot1} + \cos^{-1}\rho_{13\cdot2})/4\pi - \tfrac{1}{4}.$$

$k > 4$

It is not, in general, possible to obtain the distribution for $k > 4$ because there is no known expression for $\Pr\{z_1 > 0, z_2 > 0, \ldots, z_m > 0\}$ when $m > 3$. Some numerical values and methods have, however, been given by Plackett (1954) and these may prove useful. For the special, but important, case when the ω's are all equal, results can be obtained for $k = 5$. In this case all permutations of the x's are equally probable and so

$$\Pr(x_1 > x_2 > x_3 > x_4 > x_5) = P(5,5) = 1/5! = \tfrac{1}{120}.$$

The probabilities required to obtain $P(l, 5)$ for $l < 5$ are all known so the above methods can be used. The results obtained are given in Table 1. I have conjectured the following result for the general case with equal weights*. It yields the correct values for $k \leqslant 5$.

$$P(l,k) = \sum_{j=l-1}^{k-1} P(1, k-j)\, P(l-1, j)/l.$$

The proof is straightforward for $l = 2$ and $l = k$, but breaks down for the intermediate cases.

Table 1. *Values of $P(l, k)$ when the weights are equal*

l	$2P(l,2)$	$6P(l,3)$	$24P(l,4)$	$120P(l,5)$
1	1	2	6	24
2	1	3	11	50
3	—	1	6	35
4	—	—	1	10
5	—	—	—	1

3·2. *The distribution of $\overline{F}_k$.* The conditional distribution of $\overline{F}_k$, given that there are l means in the reduced form of the problem, is simply that of $F_{l-1,N-l}$ by the argument of § 3·1. The probabilities $P(l, k)$ depend only on the correlation coefficients ρ_{ij} which are independent of σ^2. Therefore

$$\left.\begin{aligned} \Pr\{\overline{F}_k \geqslant \gamma\} &= \sum_{2}^{k} P(l,k)\Pr\{F_{l-1,N-l} \geqslant \gamma\}, \\ \Pr\{\overline{F}_k = 0\} &= P(1,k). \end{aligned}\right\} \qquad (7)$$

For testing significance it is convenient to transform the F distribution to an incomplete B-function giving

$$\Pr\{\overline{F}_k \geqslant \gamma\} = \sum_{2}^{k} P(l,k)\,(1 - I_z\{\tfrac{1}{2}(l-1), \tfrac{1}{2}(N-l)\}) = \sum_{2}^{k} P(l,k)\, I_{1-z}(\tfrac{1}{2}(N-l), \tfrac{1}{2}(l-1)), \quad (8)$$

where $$z = (l-1)\gamma/(N-l+(l-1)\gamma).$$

* See Note added in proof on p. 48.

4. Calculation of percentage points

The results of § 3 give all that is necessary for the calculation of percentage points for the $\overline{\chi}_k^2$ and $\overline{F}_k$ statistics. Given the weights $a_1, a_2, \ldots, a_k$, and tables of the χ^2 integral or incomplete B-function, it is simply a matter of substitution in (5) or (8). The purpose of this section is to give tables of percentage points to reduce the work involved in testing significance.

Consider first the distribution of $\overline{\chi}_3^2$. From (5) we have

$$\Pr\{\overline{\chi}_3^2 \geqslant \gamma\} = \frac{1}{2}\left(1 - \frac{\cos^{-1}\rho_{12}}{\pi}\right)\int_\gamma^\infty p(\chi_2^2)\,d\chi_2^2 + \frac{1}{2}\int_\gamma^\infty p(\chi_1^2)\,d\chi_1^2.$$

Table 2 gives the values of γ satisfying

$$\Pr\{\overline{\chi}_3^2 \geqslant \gamma\} = 0{\cdot}05$$
$$= 0{\cdot}01, \quad \text{for} \quad -\rho_{12} = 0(0{\cdot}1)\,1{\cdot}0.$$

This table will be adequate for most practical work, using rough interpolation if necessary. When high accuracy is required it is better to make a direct calculation of the probability rather than use interpolation methods. This remark applies also to Table 3.

Table 2. *The* 5 *and* 1 % *points of* $\overline{\chi}_3^2$

$-\rho_{12}$	5 %	1 %	$-\rho_{12}$	5 %	1 %
0·0	4·231	7·289	0·6	3·715	6·700
0·1	4·158	7·208	0·7	3·593	6·556
0·2	4·081	7·122	0·8	3·446	6·377
0·3	4·001	7·030	0·9	3·245	6·130
0·4	3·914	6·932	1·0	2·706	5·413
0·5	3·820	6·822	—	—	—

Table 3 gives the 5 and 1 % points of $\overline{\chi}_4^2$ for $-\rho_{12}, -\rho_{23} = 0(0{\cdot}1)\,1{\cdot}0$, subject to the requirement

$$-\sqrt{(1-\rho_{23}^2)} < \rho_{12} < \sqrt{(1-\rho_{23}^2)}.$$

This restriction is simply a consequence of the fact that partial correlations must lie in the range $(-1, +1)$. As the probability integral is symmetrical in ρ_{12}, ρ_{23} the lower half only of the table is printed. Values in the upper half are obtained by interchanging ρ_{12} and ρ_{23}.

The 5 and 1 % points for equal weights when $k = 5$ are given in Table 4.

In each case the probability integral was calculated from the distribution function of χ^2 as tabulated in *Biometrika Tables for Statisticians*, and from tables of $\sin^{-1} x$. The percentage points were then obtained by inverse interpolation using the second order Bessel formula. They were checked by direct calculation using exact expressions for the χ^2 integral.

Percentage points for the $\overline{F}_k$ distribution have not yet been calculated. Although feasible for $k = 3$ a table with 3-way entry would be required for $k = 4$ and one with 4-way entry for $k = 5$. It is planned to carry out further work on the distribution of $\overline{F}_k$ to obtain, if possible, simple approximations. This would avoid the necessity for extensive tabulations.

Table 3. *The* 5 % (*upper figure*) *and* 1 % (*lower figure*) *points of* $\overline{\chi}^2_4$

$-\rho_{23}$ \ $-\rho_{12}$	0·0	0·1	0·2	0·3	0·4	0·5	0·6	0·7
0·0	5·435	—	—	—	—	—	—	—
	8·746	—	—	—	—	—	—	—
0·1	5·372	5·305	—	—	—	—	—	—
	8·676	8·600	—	—	—	—	—	—
0·2	5·307	5·235	5·160	—	—	—	—	—
	8·602	8·521	8·436	—	—	—	—	—
0·3	5·239	5·162	5·081	4·993	—	—	—	—
	8·525	8·438	8·346	8·247	—	—	—	—
0·4	5·167	5·083	4·994	4·898	4·791	—	—	—
	8·442	8·348	8·247	8·137	8·015	—	—	—
0·5	5·088	4·997	4·898	4·791	4·670	4·528	—	—
	8·351	8·248	8·137	8·014	7·873	7·709	—	—
0·6	5·003	4·900	4·789	4·665	4·524	4·354	4·135	—
	8·250	8·135	8·008	7·867	7·703	7·505	7·244	—
0·7	4·904	4·787	4·657	4·510	4·337	4·117	3·805	3·137
	8·135	8·002	7·854	7·684	7·482	7·224	6·846	6·001
0·8	4·787	4·644	4·484	4·294	4·056	3·715	2·706	—
	7·994	7·831	7·647	7·427	7·147	6·734	5·413	—
0·9	4·631	4·432	4·196	3·883	3·353	—	—	—
	7·804	7·579	7·303	6·934	6·278	—	—	—
1·0	4·231	—	—	—	—	—	—	—
	7·289	—	—	—	—	—	—	—

Table 4. *The* 5 *and* 1 % *points of* $\overline{\chi}^2_5$ *for equal weights*

5 %	1 %
5·049	8·356

5. APPLICATION OF THE TESTS

In this section three sets of data are given to which the $\overline{\chi}^2_k$ or $\overline{F}_k$ tests can be applied. The first two relate to $2 \times k$ contingency tables and are taken from the paper by Armitage (1955); the third is an artificial analysis of variance table. The justification for applying the test to $2 \times k$ contingency tables is as follows. Suppose that out of n_i observations in the ith category a proportion p_i show a certain characteristic; it is desired to test whether these proportions are equal. If each n_i is large and p_i not too close to 0 or 1 we can assume that p_i is normally distributed with variance $P(1-P)/n_i$, where P is the proportion showing the characteristic in the whole sample. The problem is thus in a form suitable for the application of the $\overline{\chi}^2_k$-test

with $a_i = n_i/\{P(1-P)\}$ $(i = 1, 2, \ldots, k)$. This approximation is, of course, the basis of the χ^2-test and is usually considered to be adequate if no expected cell frequency is less than, say, 5.

Example 1

The data shown in Table 5 are given in Armitage's paper, where references to the original work may be found. The null hypothesis is that the chance of being a carrier does not depend

Table 5. *Relationship between nasal carrier rate for streptococcus pyogenes and size of tonsils among* 1398 *children aged* 0–15 *years*

	Present but not enlarged +	Enlarged tonsils		Total
		++	+++	
Carriers	19	29	24	72
Non-carriers	497	560	269	1326
Total	516	589	293	1398
Carrier rate	0·0368	0·0492	0·0819	—

Table 6. *Changes in the size of ulcer crater,* 3 *months after treatment, for patients in two treatment groups* (*A and B*)

Treatment group	Number of cases with crater				Total
	Larger	$<\frac{2}{3}$ healed	$\geqslant\frac{2}{3}$ healed	Healed	
A	6	4	10	12	32
B	11	8	8	5	32
n_i	17	12	18	17	64
Proportion in group A	0·353	0·333	0·556	0·706	—

on size of tonsils; the alternative that the chance increases with size of tonsils. We note that the proportions are in the order expected under the alternative and hence $\overline{\chi}_3^2$ has the same value as χ^2 for this table.

Thus $\overline{\chi}_3^2 = 7{\cdot}863$,
$$-\rho_{12} = \sqrt{\left\{\frac{516 \times 293}{(516+589)(589+293)}\right\}} = 0{\cdot}394.$$

Referring to Table 2 this is found to be significant at the 1 % point in contrast to the customary χ^2, for which the significance lies between $2\frac{1}{2}$ and 1 %.

Example 2

Table 6, also from Armitage's paper, gives the results of an experiment to compare two methods of ulcer treatment.

We test the null hypothesis that the two treatments are equally good, against the alternative that treatment A is better than treatment B. Under this alternative we should expect

the proportions to increase as we move from left to right across the table. In this example the proportions are not in the expected order to begin with. Because the first is greater than the second, the first two groups are combined into one giving a table with three groups. This new table has the proportions in the expected order. $\overline{\chi}^2_4$ is obtained by calculating χ^2 for this pooled table, giving

$$\overline{\chi}^2_4 = 5{\cdot}900, \quad -\rho_{12} = 0{\cdot}593, \quad -\rho_{23} = 0{\cdot}441.$$

Extracting the relevant part of Table 3, remembering that ρ_{12} and ρ_{23} can be interchanged, we have for the 5 % points,

$-\rho_{23}$ \ $-\rho_{12}$	0·5	0·6
0·4	4·670	4·524
0·5	4·528	4·354

$\overline{\chi}^2_4$ is therefore clearly significant at the 5 % level (but not at 1 %). If the χ^2-test had been used the significance would have been between 11 and 12 %.

Example 3

The following table gives artificial data constructed to demonstrate the application of the $\overline{F}_k$-test to an analysis of variance table.

Table 7. *Artificial data to illustrate the application of the $\overline{F}_k$ test*

Group ...	I	II	III	IV	V	Total
Total	110	123	126	140	96	595
No. observed	4	4	4	4	4	20
Sum of squares	3,380	3,601	3,712	3,984	3,201	12,378
Mean	27·50	30·75	31·50	35·00	24·00	—
				29·50		
			30·17			
		30·31				

It is supposed that the alternative predicts that the means will increase as we move from left to right across the table. We begin by averaging the means of groups IV and V followed by III, IV and V and finally II, III, IV and V as indicated underneath the table. The reduced form of the problem therefore consists of two groups; group I with mean 27·50 and groups II–V with mean 30·31. Calculating the F ratio for the reduced problem we obtain

$$F = 3{\cdot}008,$$

which is the value of $\overline{F}_k$ for Table 7. To test the significance we require, from (8)

$$z = \frac{3{\cdot}008}{3{\cdot}008 + 18} = 0{\cdot}1432.$$

The following table sets out the method of calculating $\Pr\{\overline{F}_5 \geqslant 3{\cdot}008\}$. As the numbers in each group are equal the probabilities $P(l, k)$ are obtained directly from Table 1.

There is thus no reason to reject the null hypothesis that the groups are drawn from populations with the same mean.

Table 8. *Calculations required for testing the significance of* $\overline{F}_k$

l	$P(l, k)$	$I_{1-z}(\frac{1}{2}(l-1), \frac{1}{2}(N-l))$	
		$z = 0{\cdot}14$	$z = 0{\cdot}15$
2	50/120	0·1041	0·0916
3	35/120	0·2775	0·2512
4	10/120	0·4779	0·4439
5	1/120	0·6614	0·6281

$\Pr(z \geqslant 0{\cdot}14) = 0{\cdot}1696$, $\Pr(z \geqslant 0{\cdot}15) = 0{\cdot}1537$.

6. Further work

As pointed out at the end of § 4 the greatest need is for tables and approximate methods to simplify the method for testing the significance of $\overline{F}_k$. It would also be an advantage to obtain values of $P(l, k)$ for $k \geqslant 5$. If the conjecture given in § 3 above can be proved, then the distribution will be completely determined for the case of equal weights. Until this problem has been solved the best course appears to be the reduction of the number of groups by pooling. This will undoubtedly result in a loss of efficiency but even so may be better than the usual χ^2-test, which takes no account of the order. The answer to this question would require the calculation of the power function of $\overline{\chi}^2_k$. A new problem arises in the analysis of data like that of example 2 in § 4. When applying the $\overline{\chi}^2_k$-test it was assumed that treatment A would be the better of the two treatments. If we had wished to test for a difference without specifying which would be the better the alternative would have had to be,

$$H_3: m_1 \geqslant m_2 \geqslant \ldots \geqslant m_k,$$
$$\text{or} \quad m_1 \leqslant m_2 \leqslant \ldots \leqslant m_k.$$

It is hoped to communicate the results of work on these topics later.

I am indebted to the referee for suggesting a shorter method of deriving the test criterion and for pointing out a gap in the original derivation of equation (5).

Editorial Note added in proof. The recurrence relation for the probabilities $P(l, k)$ conjectured in §3.1 has now been proved by Mr R. E. Miles. He has also established the more useful recurrence relation

$$P(l, k) = [P(l-1, k-1) + (k-1)\,P(l, k-1)]/k$$

for $2 \leqslant l \leqslant k-1$, which, together with $P(1, k) = k^{-1}$ and $P(k, k) = (k!)^{-1}$, determine the probabilities easily. It is hoped to publish the proof in the next issue of *Biometrika*. This result can now be used for testing the significance of $\overline{\chi}^2_k$ and $\overline{F}_k$ for $k > 5$ if the weights are equal.

REFERENCES

Armitage, P. (1955). Tests for linear trends in proportions and frequencies. *Biometrics*, **11**, 375–86.
David, F. N. (1953). A note on the evaluation of the multivariate normal integral. *Biometrika*, **40**, 458–9.
Jonckheere, A. R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika*, **41**, 133–45.
Plackett, R. L. (1954). A reduction formula for normal multivariate integrals. *Biometrika*, **41**, 351–60.

SOME PROBLEMS INVOLVING LINEAR HYPOTHESES IN MULTIVARIATE ANALYSIS

By C. RADHAKRISHNA RAO
Indian Statistical Institute, Calcutta

1. Introduction

We consider p correlated normal variables $y_1, \ldots, y_p$ with the following specification.

(i) $$E(y_i) = a_{i1}\tau_1 + \ldots + a_{im}\tau_m, \tag{1·1}$$

or in matrix notation $E(\mathbf{y}) = \mathbf{A}\boldsymbol{\tau}$, where $\mathbf{y}$ is the column vector of the random variables $y_1, \ldots, y_p$; $\boldsymbol{\tau}$ that of m unknown parameters $\tau_1, \ldots, \tau_m$ and $\mathbf{A}$, the matrix of coefficients is known.

(ii) The dispersion matrix of $y_1, \ldots, y_p$ is $e^{-1}\boldsymbol{\Lambda}$ where e is a known constant and $\boldsymbol{\Lambda}$ is unknown, but an estimate $\mathbf{S}$ of $\boldsymbol{\Lambda}$ is available and has Wishart's distribution based on f degrees of freedom independently of $y_1, \ldots, y_p$. The matrix $\boldsymbol{\Lambda}$ is assumed to be non-singular.

The set-up is thus seen to be that of the theory of least squares with correlated variables generalized to the case where the variances and covariances are unknown, but an estimate is available. The estimates of parameters as well as the test criteria suggested in this paper are essentially those derived by the theory of least squares, substituting the estimate $\mathbf{S}$ for the unknown $\boldsymbol{\Lambda}$. Fortunately, the distribution problems present no difficulty and valid inferences can be drawn with the help of the existing t and F tables only.

A simple case where such a set-up holds is when we have a sample of size n from a p-variate normal distribution with an unknown dispersion matrix $\boldsymbol{\Lambda}$ and the mean of x_i, the ith variable, is as specified in (1·1). Here

$$y_i = \bar{x}_i, \quad e = \frac{1}{n} \tag{1·2}$$

and $\mathbf{S}$ is the matrix of estimated variances and covariances on $f = (n-1)$ degrees of freedom.

In the general set-up, on the basis of a given realization of $y_1, \ldots, y_p$ and $\mathbf{S}$ we wish to examine the following problems.

(i) Is the specification (1·1) adequate? Thus, if $y_1, \ldots, y_p$ represent the average weights of a group of individuals at p time points, we may ask whether the average growth curve can be represented by a polynomial in time of a given degree. In such a case the parameters represent the coefficients of the polynomial trend.

(ii) If (1·1) is true, how can estimates of $\tau_1, \ldots, \tau_m$ be obtained and the precision of estimates expressed?

(iii) Assuming the specification (1·1), how can general linear hypotheses concerning $\tau_1, \ldots, \tau_m$ be tested?

(iv) How can simultaneous confidence limits to a class of linear functions of $\tau_1, \ldots, \tau_m$ be obtained?

Problem (i) is a generalization of some particular cases considered earlier, such as testing the equality of means of p-correlated variables, equality of means of subsets of variables, etc. (Hsu, 1938; Rao, 1946*b*, 1952). Problem (iv), which is the main motivation for the present

paper, is an attempt to obtain confidence bands of a smaller width (see § 4) than those given by Hoel (1954) in the case of linear regression such as polynomial growth curves, dosage-response relationships, etc., which are covered by the general set-up (1·1).

2. Test for adequacy of the specification (1·1)

Let $\mathbf{B}$ be a matrix with r, equal to p minus the rank of $\mathbf{A}$, independent vectors

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{21} & \dots & b_{p1} \\ . & . & \dots & . \\ b_{1r} & b_{2r} & \dots & b_{pr} \end{pmatrix} \tag{2·1}$$

such that the vectors in $\mathbf{B}$ are orthogonal to those in $\mathbf{A}'$. The computation of $\mathbf{B}$, known as the deficiency matrix of $\mathbf{A}'$, can be conveniently carried out by the method of sweep out (Rao, 1952, § 1 a·3). It is easy to see that the variables $z_1, \dots, z_r$ defined by

$$\left.\begin{aligned} z_1 &= b_{11}y_1 + \dots + b_{p1}y_p, \\ . \quad & \quad . \quad \dots \quad . \\ z_r &= b_{1r}y_1 + \dots + b_{pr}y_p, \end{aligned}\right\} \tag{2·2}$$

are such that

$$E(z_i) = 0, \quad (i = 1, \dots, r) \tag{2·3}$$

independently of the unknown parameters $\boldsymbol{\tau}$, and with dispersion matrix equal to $e^{-1}\mathbf{B}\boldsymbol{\Lambda}\mathbf{B}'$. Given $y_1, \dots, y_p$ and $\mathbf{S}$ we can compute $z_1, \dots, z_r$ by (2·2) and their estimated dispersion matrix $e^{-1}\mathbf{BSB}' = e^{-1}\mathbf{G}$ (say). It is easy to see that $\mathbf{G}$ itself has a Wishart distribution based on the same degrees of freedom f as that of $\mathbf{S}$ and is distributed independently of $z_1, \dots, z_r$.

The problem of testing the adequacy of the specification (1·1) is the same as that of testing whether the r correlated variables defined in (2·2) have zero mean values. The appropriate statistic for this is Hotelling's T defined by

$$T_r = \frac{e}{f}\mathbf{z}'\mathbf{G}^{-1}\mathbf{z} = \frac{e}{f}\sum_{ij} g^{ij} z_i z_j \tag{2·4}$$

with the corresponding variance ratio (Rao, 1952, p. 71)

$$\frac{f+1-r}{r} T_r \tag{2·5}$$

based on r and $(f+1-r)$ degrees of freedom. The matrix $\mathbf{B}$ is not unique but the statistic T_r is invariant as we shall see below.

An alternative way of deriving the test criterion is to consider first the statistic appropriate to test the specification (1·1) when $\boldsymbol{\Lambda}$ is known. This according to the least squares theory is the residual quadratic form

$$\min e(\mathbf{y} - \mathbf{A}\boldsymbol{\tau})' \boldsymbol{\Lambda}^{-1}(\mathbf{y} - \mathbf{A}\boldsymbol{\tau}) \tag{2·6}$$

obtained by fitting the parameters by the method of least squares. This residual quadratic form (2·6) for known $\boldsymbol{\Lambda}$ is distributed as χ^2 on r degrees of freedom. When $\boldsymbol{\Lambda}$ is unknown we may consider the statistic

$$\min e(\mathbf{y} - \mathbf{A}\boldsymbol{\tau})' \mathbf{S}^{-1}(\mathbf{y} - \mathbf{A}\boldsymbol{\tau}) \tag{2·7}$$

obtained by substituting the estimate $\mathbf{S}$ for $\boldsymbol{\Lambda}$ in (2·6).

Given the matrix $\mathbf{B}$ of (2·1) it is easy to see that there exists a matrix $\mathbf{F}$ of rank $(p-r)$ and order $(p-r, r)$ such that $\mathbf{BSF}' = 0$, and with such a choice of F (Rao, 1952, p. 53)

$$(\mathbf{y} - \mathbf{A}\boldsymbol{\tau})' \mathbf{S}^{-1}(\mathbf{y} - \mathbf{A}\boldsymbol{\tau}) = (\mathbf{By})' (\mathbf{BSB}')^{-1}(\mathbf{By}) + (\mathbf{Fy} - \mathbf{FA}\boldsymbol{\tau})'(\mathbf{FSF}')^{-1}(\mathbf{Fy} - \mathbf{FA}\boldsymbol{\tau}) \tag{2·8}$$

which is the sum of two positive definite quadratic forms with the first expression independent of $\boldsymbol{\tau}$. By choosing $\boldsymbol{\tau}$ such that $\mathbf{Fy} = \mathbf{FA}\boldsymbol{\tau}$, the second expression becomes zero, leading to the result

$$\min (\mathbf{y} - \mathbf{A}\boldsymbol{\tau})' \mathbf{S}^{-1} (\mathbf{y} - \mathbf{A}\boldsymbol{\tau}) = (\mathbf{By})' (\mathbf{BSB}')^{-1} (\mathbf{By}) = \mathbf{z}'\mathbf{G}^{-1}\mathbf{z}. \tag{2·9}$$

Thus the statistic (2·7) is same as T_r defined in (2·4) apart from a constant multiplier. Incidentally the invariance of T_r with respect to the choice of $\mathbf{B}$ is also established.

We, thus, have two methods of computing the statistic T_r either to find a $\mathbf{B}$ and compute the quadratic form $\mathbf{z}'\mathbf{G}^{-1}\mathbf{z}$ or obtain the minimum value in (2·6) directly by some other method.

3. Tests of linear hypotheses

In this section we consider the problem of testing linear hypotheses assuming the specification (1·1) to be true. Let

$$\left.\begin{aligned} u_{11}\tau_1 + \ldots + u_{1m}\tau_m &= \xi_1, \\ \cdot \quad \cdot \quad \ldots \quad \cdot & \\ u_{k1}\tau_1 + \ldots + u_{km}\tau_m &= \xi_k, \end{aligned}\right\} \tag{3·1}$$

or in matrix notation $\mathbf{U}\boldsymbol{\tau} = \boldsymbol{\xi}$, where $\mathbf{U}$ and $\boldsymbol{\xi}$ are known, be a set of linear hypotheses to be tested. The rank of $\mathbf{U}$ is taken to be k. Let us suppose that there exist linear functions $z_{r+1}, \ldots, z_{r+k}$ of $y_1, \ldots, y_p$ such that

$$E(z_{r+i}) = u_{i1}\tau_1 + \ldots + u_{im}\tau_m. \quad (i = 1, \ldots, k) \tag{3·2}$$

A linear function $u_1\tau_1 + \ldots + u_m\tau_m$ is said to be not estimable if it cannot be expressed as an expectation of a linear function of the variables $y_1, \ldots, y_p$. A linear hypothesis cannot be tested when the parametric function associated with it is not estimable, since such a parametric function can assume any arbitrary value without altering the probability of the observations. In the case of a set of linear hypotheses such as (3·1) each of the associated linear functions must be estimable (Rao, 1952, p. 83), the condition for which is that there exists a matrix $\mathbf{L}$ of order (k, p) such that $\mathbf{U} = \mathbf{LA}$, where $\mathbf{A}$ is the matrix of (1·1).

The hypothesis (3·1) together with the specification (1·1) is equivalent to the hypothesis

$$E(z_{r+1}) = \xi_1, \ldots, E(z_{r+k}) = \xi_k \tag{3·3}$$

given that

$$E(z_1) = 0, \ldots, E(z_r) = 0, \tag{3·4}$$

where $z_1, \ldots, z_r$ are as defined in (2·2). Independent estimates of the variances and covariances of $z_1, \ldots, z_{r+k}$ derivable from $\mathbf{S}$ have a Wishart distribution on the same degrees of freedom. The problem of testing a hypothesis such as (3·3) given (3·4) was discussed in earlier papers (Rao, 1946*b*, 1949). The appropriate test is as follows.

Let $\mathbf{D}$ be the transformation matrix defining $z_1, \ldots, z_{r+k}$ in terms of $y_1, \ldots, y_p$. The estimated variance-covariance matrix of $z_1, \ldots, z_{r+k}$ is $\mathbf{DSD}' = \mathbf{H}$ (say). If $\mathbf{H}^{-1} = (h^{ij})$ is the reciprocal matrix, then the statistic for testing given values of expectations of all the $(r+k)$ variables is

$$T_{r+k} = \frac{e}{f} \sum_{i,j=1}^{r+k} h^{ij} [z_i - E(z_i)] [z_j - E(z_j)]. \tag{3·5}$$

The statistic T_r for testing the expectations (3·4) is already defined in (2·4). The statistic for testing the hypothesis (3·3) given (3·4) is

$$\frac{f+1-k-r}{k} \frac{T_{r+k} - T_r}{1+T_r}, \tag{3·6}$$

which is a variance ratio based on k and $(f+1-k-r)$ degrees of freedom (Rao, 1952, p. 83). We observe that the value of (3·6) would be the same for all choices of linear functions $z_{r+1}, \ldots, z_{r+k}$ such that their expectations are as in (3·2). The optimum properties of a test of the type (3·6) is discussed in an earlier paper (Rao, 1949).

An alternative method of computing T_{r+s}, without necessarily making a transformation to new variables, is provided by the following equation

$$\frac{f}{e}T_{r+k} = \min\,(\mathbf{y}-\mathbf{A}\boldsymbol{\tau})'\,\mathbf{S}^{-1}(\mathbf{y}-\mathbf{A}\boldsymbol{\tau}), \tag{3·7}$$

where the minimization is subject to conditions (3·1) on $\boldsymbol{\tau}$; whereas

$$\frac{f}{e}T_{r} = \min\,(\mathbf{y}-\mathbf{A}\boldsymbol{\tau})'\,\mathbf{S}^{-1}(\mathbf{y}-\mathbf{A}\boldsymbol{\tau}), \tag{3·8}$$

the minimization being unrestricted.

A third method of computing T_{r+k} or directly the difference $T_{r+k}-T_r$ is provided by the theory of least squares (Rao, 1946*a*). This makes use of the estimates of $\boldsymbol{\tau}$ obtained by minimizing, unconditionally, $(\mathbf{y}-\mathbf{A}\boldsymbol{\tau})'\,\mathbf{S}^{-1}(\mathbf{y}-\mathbf{A}\boldsymbol{\tau})$ for which the normal equations are

$$\mathbf{A}'\mathbf{S}^{-1}\mathbf{y} = \mathbf{A}'\mathbf{S}^{-1}\mathbf{A}\hat{\boldsymbol{\tau}}, \tag{3·9}$$

where $\hat{\boldsymbol{\tau}}$ represents the estimates. Let $\mathbf{C}$ denote the reciprocal of $\mathbf{A}'\mathbf{S}^{-1}\mathbf{A}$, if it exists, or a pseudo-inverse as defined by Rao (1955), and

$$d_i = u_{i1}\hat{\tau}_1 + \ldots + u_{im}\hat{\tau}_m \quad (i = 1, \ldots, k). \tag{3·10}$$

Then
$$\frac{f}{e}(T_{r+k}-T_r) = (\mathbf{d}-\boldsymbol{\xi})'\,(\mathbf{UCU}')^{-1}(\mathbf{d}-\boldsymbol{\xi}) \tag{3·11}$$

where $\mathbf{U}$ and $\boldsymbol{\xi}$ are given in (3·1). Substituting for the numerator in (3·6) the right-hand side expression of (3·11) we find that

$$\frac{(f+1-k-r)e}{kf}\,\frac{(\mathbf{d}-\boldsymbol{\xi})'\,(\mathbf{UCU}')^{-1}(\mathbf{d}-\boldsymbol{\xi})}{1+T_r} \tag{3·12}$$

has a variance ratio distribution on k and $(f+1-k-r)$ degrees of freedom.

4. SIMULTANEOUS CONFIDENCE LIMITS TO A CLASS OF LINEAR PARAMETRIC FUNCTIONS

Let us first consider the problem of setting up confidence limits to a single estimable parametric function

$$\lambda = u_1\tau_1 + \ldots + u_m\tau_m. \tag{4·1}$$

Using the notation adopted in § 2, we find from (3·12) that if $d = u_1\hat{\tau}_1 + \ldots + u_m\hat{\tau}_m$ and $g = \mathbf{u}'\mathbf{Cu}$

$$(d-\lambda)\sqrt{\left(\frac{e(f-r)}{gf(1+T_r)}\right)} \tag{4·2}$$

has a t-distribution on $(f-r)$ degrees of freedom. Hence, the confidence limits to λ are

$$d \pm t\left(\frac{(f-r)\,e}{gf(1+T_r)}\right)^{-\frac{1}{2}}. \tag{4·3}$$

In particular, confidence limits to any single parameter τ_i can be obtained by choosing $\lambda = \tau_i$ provided it is estimable. The confidence limits to τ_i are

$$\hat{\tau}_i \pm t\left(\frac{e(f-r)}{c_{ii}f(1+T_r')}\right)^{-\frac{1}{2}}, \tag{4·4}$$

where c_{ii} is the ith diagonal element in the matrix $\mathbf{C}$.

Now we let the vector $\mathbf{u}$ belong to the $k(\leqslant m)$ dimensional subspace spanned by the vectors

$$\mathbf{U} = \begin{pmatrix} u_{11} & \dots & u_{1m} \\ . & \dots & . \\ u_{k1} & \dots & u_{km} \end{pmatrix} \tag{4·5}$$

and represent, as in (3·10),

$$d_i - \lambda_i = (u_{i1}\hat{\tau}_1 + \dots + u_{im}\hat{\tau}_m) - \lambda_i. \quad (i = 1, \dots, k). \tag{4·6}$$

Then it is known that (Rao, 1946*a*, 1952, p. 85),

$$\max_{\mathbf{u}} \frac{(\mathbf{u}'\hat{\boldsymbol{\tau}} - \mathbf{u}'\boldsymbol{\tau})^2}{\mathbf{u}'\mathbf{C}\mathbf{u}} = (\mathbf{d}-\boldsymbol{\lambda})'(\mathbf{U}\mathbf{C}\mathbf{U}')^{-1}(\mathbf{d}-\boldsymbol{\lambda}) \tag{4·7}$$

when $\mathbf{u}$ is restricted to the subspace (4·5). Using the distribution result (3·12)

$$\max_{\mathbf{u}} \frac{(\mathbf{u}'\hat{\boldsymbol{\tau}} - \mathbf{u}'\boldsymbol{\tau})^2}{\mathbf{u}'\mathbf{C}\mathbf{u}} \frac{(f+1-k-r)\,e}{fk(1+T_r')} \tag{4·8}$$

has an F-distribution with k and $(f+1-k-r)$ degrees of freedom. If F_α represents the α-value of F corresponding to these degrees of freedom then the probability that

$$\frac{(\mathbf{u}'\hat{\boldsymbol{\tau}} - \mathbf{u}'\boldsymbol{\tau})^2}{\mathbf{u}'\mathbf{C}\mathbf{u}} \leqslant F_\alpha \frac{kf(1+T_r')}{e(f+1-k-r)} \tag{4·9}$$

for all $\mathbf{u}$ in the subspace spanned by the vectors in (4.5) is not less than $1-\alpha$. The simultaneous confidence limits for all $\mathbf{u}'\boldsymbol{\tau}$ where $\mathbf{u}'$ belongs to $\mathbf{U}$ is thus provided by

$$u_1\hat{\tau}_1 + \dots + u_m\hat{\tau}_m \pm \sqrt{\left(F_\alpha \frac{kf(1+T_r')}{e(f+1-k-r)}\right)}. \tag{4·10}$$

The argument employed here is the same as that used by Tukey (1951), Scheffé (1953), Roy & Bose (1953), Hoel (1954) and others. It may be observed that the first attempt to obtain such simultaneous confidence intervals was due to Hotelling & Working (1929). The result (4·10) differs from that given by Hoel (1954) in that, in effect, he bases the confidence interval on the distribution of F on p, the number of variables $\mathbf{y}$, and $(f+1-p)$ degrees of freedom irrespective of the dimensionality of the subspace of vectors $\mathbf{u}'$ for which simultaneous confidence intervals are intended. Unless the factor $(1+T_r')$ in (4·10) is very large, which may happen when the specification (1·1) is not correct, in which case the procedure suggested is not valid, the confidence interval based on F, with k and $(f+1-k-r)$ degrees of freedom should be smaller in length when $k < p$.

5. AN ILLUSTRATIVE EXAMPLE

The results of the preceding sections show that the computations involved in tests of hypotheses and estimation of parameters are the same as in the method of least squares with correlated variables, although the distributions are slightly different due to the estimated dispersion matrix being used in the analysis instead of the true one. The computations can be

carried up to a certain stage mechanically as in the method of least squares with correlated variables. We may briefly review the methods to make this section self-contained for the use of those interested in the computational techniques.

Consider the quadratic form

$$(\mathbf{y}-\mathbf{A}\boldsymbol{\tau})'\,\mathbf{S}^{-1}(\mathbf{y}-\mathbf{A}\boldsymbol{\tau}) \tag{5·1}$$

with the estimated matrix $\mathbf{S}$ for the unknown dispersion matrix and obtain the normal equations

$$\mathbf{A}'\mathbf{S}^{-1}\mathbf{y} = \mathbf{A}'\mathbf{S}^{-1}\mathbf{A}\hat{\boldsymbol{\tau}}. \tag{5·2}$$

Find a solution $\hat{\boldsymbol{\tau}}$ and a matrix $\mathbf{C}$ of the variances and covariances of the estimates according to least squares theory assuming $\mathbf{S}$ to be true. The matrix $\mathbf{C}$ is the inverse of the matrix $\mathbf{A}'\mathbf{S}^{-1}\mathbf{A}$, and is unique if the rank of $\mathbf{A}$ is equal to the number of unknown parameters or a pseudo-inverse (which need not be unique) otherwise. The minimum value of the expression (5·1) is given by

$$\mathbf{y}'\mathbf{S}^{-1}\mathbf{y}-\mathbf{y}'\mathbf{S}^{-1}\mathbf{A}\hat{\boldsymbol{\tau}}. \tag{5·3}$$

The minimum value (5·3), $\hat{\boldsymbol{\tau}}$, a solution of the normal equations, and $\mathbf{C}$, which will be referred to as a least squares variance-covariance matrix, are used in obtaining confidence intervals and testing hypotheses as shown in the following illustrative example.

The example refers to observations on growth of ten rats under a control treatment considered in an earlier paper (Rao, 1958). The averages of 10 values representing increases in growth during 4 successive weeks are

$$y_0 = 24{\cdot}5, \quad y_1 = 27{\cdot}5, \quad y_2 = 24{\cdot}1, \quad y_3 = 30{\cdot}5$$

and the estimated dispersion matrix $\mathbf{S}$, within treatments, based on 24 D.F. (pooled from other treatment groups also) is

24·2625	1·7708	−2·3125	−3·1083
	25·3750	26·1042	14·3542
		43·6125	19·1250
			35·5417

In the notation of § 1, $e = 10$ and $f = 24$. Let us examine whether a second degree polynomial in time can be fitted to the first differences of the growth curve or in other words whether

$$E(y_t) = \alpha+\beta t+\gamma t^2 \quad (t = 0, 1, 2, 3), \tag{5·4}$$

so that $r = 4-3 = 1$. The unknown parameters are α, β and γ and the matrix $\mathbf{A}$ of coefficients is

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$$

corresponding to the expectations for $t = 0, 1, 2, 3$.

The first step is to obtain the normal equations

$$\mathbf{A}'\mathbf{S}^{-1}\mathbf{y} = \mathbf{A}'\mathbf{S}^{-1}\mathbf{A}\hat{\boldsymbol{\tau}}$$

which involves the evaluation of the inverse of $\mathbf{S}$ and a triple product of matrices. A simple way of doing this is first to obtain a triangular reduction of the matrix $\mathbf{S}$, with the vector $\mathbf{y}$ and matrix $\mathbf{A}$ appended, by the square root or Gauss–Doolittle method (Dwyer, 1951; Rao, 1955).

Table 1. *Triangular reduction of matrix* **S**, *with* **y** *and* **A** *appended, for obtaining normal equations (square root method)*

Row no.	Matrix S				y	Matrix A		
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0·1	24·2625	1·7708	−2·3125	−3·1083	24·5	1	0	0
0·2	—	25·3750	26·1042	14·3542	27·5	1	1	1
0·3	—	—	43·6125	19·1250	24·1	1	2	4
0·4	—	—	—	35·5417	30·5	1	3	9
1·1	4·925698	0·359502	−0·469477	0·631037	4·973914	0·203017	0	0
1·2	—	5·024516	5·228957	2·901983	5·117282	0·184498	0·199024	0·199024
1·3	—	—	4·006257	0·912171	−0·080604	0·032591	0·239451	0·738670
1·4	—	—	—	5·088214	3·706994	0·110642	0·433160	1·522860

The rows (0·1) to (0·4) contain the matrix **S**, the vector **y** and the matrix **A**. We consider the *entire matrix* which is now in 8 columns and rows (0·1) to (0·4) and reduce it step by step by the square root method till the matrix **S** is reduced to the diagonal form. In this case only 4 steps are needed. The figures in columns (5) to (8) in the rows (1·1) to (1·4) containing the reduced matrices are used for further computations. We define the product of two columns as the sum of products of corresponding elements taken from rows (1·1) to (1·4). Thus

$$[5,7] = (4{\cdot}973914)(0) + (5{\cdot}117282)(0{\cdot}199024) + (-0{\cdot}080604)(0{\cdot}239451) + (3{\cdot}706994)(0{\cdot}433160) = 2{\cdot}604883,$$

$$[5,5] = (4{\cdot}973914)^2 + \ldots + (3{\cdot}706994)^2 = 64{\cdot}674697.$$

The normal equations are

$$\alpha[6,6] + \beta[6,7] + \gamma[6,8] = [6,5],$$
$$\alpha[7,6] + \beta[7,7] + \gamma[7,8] = [7,5],$$
$$\alpha[8,6] + \beta[8,7] + \gamma[8,8] = [8,5].$$

Computing these products and writing the equations in a form suitable for further computations we have

α	β	γ	Q
0·088560	0·092449	0·229286	2·361459
—	0·284575	0·876129	2·604883
—	—	2·904349	6·604173

where **Q** with elements $Q_\alpha, Q_\beta, Q_\gamma$, stands for the products of column (5) containing the vector **y** with the others. The reciprocal of the left-hand side matrix is

$$\mathbf{C} = \begin{pmatrix} 22{\cdot}867188 & -26{\cdot}251429 & 6{\cdot}113767 \\ & 79{\cdot}444013 & -21{\cdot}892726 \\ & & 6{\cdot}465838 \end{pmatrix}$$

and by multiplying the rows of this matrix with the column **Q** we obtain estimates of α, β, γ as

$$\hat{\alpha} = 25{\cdot}994127, \quad \hat{\beta} = 0{\cdot}368153, \quad \hat{\gamma} = 0{\cdot}110708.$$

The elements of the matrix **C** represent the variances and covariances of the estimates only when **S** is the true dispersion matrix. They have to be used slightly differently in expressing the precision of the estimates in the present case. The minimum value of $(\mathbf{y}-\mathbf{A}\boldsymbol{\tau})'\,\mathbf{S}^{-1}(\mathbf{y}-\mathbf{A}\boldsymbol{\tau})$ is $[5,5]-\hat{\alpha}Q_\alpha-\hat{\beta}Q_\beta-\hat{\gamma}Q_\gamma = 1{\cdot}6005$. If the object is to obtain this minimum value only then it is not necessary to obtain explicitly the solutions of the normal equations. The triangular reduction of the normal equations, say by the square root method, gives three values in the column for **Q**, the sum of squares of which is equal to the expression $\hat{\alpha}Q_\alpha+\hat{\beta}Q_\beta+\hat{\gamma}Q_\gamma$. If estimates of parameters and their precisions are also required the method suggested above may be followed.

Test for specification. In the present problem, it is specified that the average differential growth curve of the population of rats is a second degree polynomial in time. To test whether this specification is adequate or not, we use the statistic (2·5, 2·9)

$$\frac{(f+1-r)e}{rf}\min_{\boldsymbol{\tau}}(\mathbf{y}-\mathbf{A}\boldsymbol{\tau})'\,\mathbf{S}^{-1}(\mathbf{y}-\mathbf{A}\boldsymbol{\tau}),$$

which is a variance ratio with degrees of freedom $r = 1$ and $(f+1-r) = 24$ in the present problem. The value of the statistic

$$\tfrac{24}{1}\tfrac{10}{24}(1{\cdot}6005) = 16{\cdot}005 \tag{5·5}$$

exceeds 4·25, the 5 % point of F on 1 and 24 D.F., thus indicating inadequacy of the specification at that level of probability.

It may be observed that if the object is only to test the adequacy of specification (5·4), the alternative method of computing the matrix **B** and the quadratic form (2·4) suggested in §2 would be simpler in the present case. The matrix $\mathbf{A}'$ is

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix}$$

and by inspection we observe that the matrix with rank $r = 1$, which is orthogonal to $\mathbf{A}'$ is

$$\mathbf{B} = (-1 \quad 3 \quad -3 \quad 1),$$

from which we compute

$$z_1 = (-1)y_0+3y_1-3y_2+y_3 = 16{\cdot}2,$$
$$\mathbf{BSB}' = 163{\cdot}9081.$$

Since there is only one linear function z_1, the appropriate statistic is Student's t. We can refer

$$t^2 = \frac{f+1-1}{1}T_1 = \frac{f}{1}\frac{e}{f}\frac{z_1^2}{\mathbf{BSB}'} = 16{\cdot}011 \tag{5·6}$$

to an F table with 1 and 24 D.F. This value agrees, to the degree of approximation expected, with that in (5·5) obtained by a different procedure. Although the second method is more direct, it does not provide estimates of the constants, which may be needed in actual practice. But if the sole object is a test for adequacy of the specification then the second method may be simpler.

Precision of estimates. In view of the fact that the assumed specification is not found to be adequate by the above test the problem of providing estimates of the parameters α, β, γ and

their precisions does not arise. We need to look for an alternative specification which will be supported by the data. But for purposes of illustration the computations laid out above have been used in the following, bearing in mind that the estimates and their precisions acquire a meaning only when the assigned specification is not rejected by a test.

The confidence interval for a parametric function $u_1\tau_1+\ldots+u_m\tau_m$ is given by the formula (4·3). Obtain the least squares estimate

$$d = u_1\hat{\tau}_1+\ldots+u_m\hat{\tau}_m$$

and least squares variance, $g = \mathbf{u'Cu}$. The confidence interval is

$$d \pm t\left(\frac{(f-r)\,e}{gf(1+T_r)}\right)^{-\frac{1}{2}},$$

where

$$T_r = \frac{e}{f}\min\,(\mathbf{y}-\mathbf{A\tau})'\,\mathbf{S}^{-1}(\mathbf{y}-\mathbf{A\tau}).$$

In the present example, if a confidence interval for the parameter α is desired the formula is

$$\hat{\alpha} \pm t\left(\frac{(f-r)\,e}{c_{11}f(1+T_r)}\right)^{-\frac{1}{2}},$$

where c_{11}, the first diagonal element in the matrix $\mathbf{C}$, is the least squares variance of the estimate of α. Substituting the observed values we have the limits $25{\cdot}9941 \pm 4{\cdot}1263$, corresponding to 95% probability.

Confidence bands to the differential growth curve. Assuming the specification to be true, let us consider the problem of obtaining confidence bands to the polynomial $\alpha+\beta t+\gamma t^2$ (for all values of the time). Although vectors of the form $(1, t, t^2)$ do not cover all the points of a three-dimensional space the probability statement (4·9) applicable to all combinations of α, β, γ is still true as an inequality. For an arbitrary value of time t the estimated polynomial value is

$$\alpha+\beta t+\gamma t^2 = 25{\cdot}9941+0{\cdot}3682t+0{\cdot}1107t^2$$

and the least squares variance of the estimate is the product

$$(1,t,t^2)\,\mathbf{C}\begin{pmatrix}1\\t\\t^2\end{pmatrix} = 22{\cdot}8672-52{\cdot}5029t+91{\cdot}6715t^2-43{\cdot}7855t^3+6{\cdot}4658t^4.$$

The other values needed for an application of (4·10) are

$$k=3,\quad f=24,\quad e=10,\quad 1+T_r=1{\cdot}6669,\quad F_\alpha=3{\cdot}07,\quad \alpha=0{\cdot}05.$$

Substituting $t = 0, 1, 2, 3$ we obtain the following confidence limits:

	$t=0$	$t=1$	$t=2$	$t=3$
Lower limit	19·6600	19·8877	19·0383	20·6107
Mean value	25·9941	26·4729	27·1731	28·0947
Upper limit	32·3282	33·0581	35·3079	35·5787

From these values the limits to other values of t can be obtained by graduation (such as free hand smoothing) if necessary. The confidence bands are valid only when the specification adopted is supported by the data, for which a preliminary test is necessary.

REFERENCES

DWYER, P. S. (1951). *Linear Computations.* New York: John Wiley and Sons.

HOEL, P. G. (1954). Confidence bands for polynomial curves. *Ann. Math. Statist.* **25**, 534.

HOTELLING, H. & WORKING, H. (1929). Applications of the theory of errors to the interpretation of trends. *J. Amer. Statist. Ass.* (Suppl.), **24**, 73.

HSU, P. L. (1938). Notes on Hotelling's generalized T. *Ann. Math. Statist.* **9**, 231.

RAO, C. R. (1946*a*). On the linear combination of observations and the general theory of least squares. *Sankhyā*, **7**, 237.

RAO, C. R. (1946*b*). Tests with discriminant functions in multivariate analysis. *Sankhyā*, **7**, 407.

RAO, C. R. (1949). On some problems arising out of discrimination with multiple characters. *Sankhyā*, **9**, 343.

RAO, C. R. (1952). *Advanced Statistical Methods in Biometric Research.* New York: John Wiley and Sons.

RAO, C. R. (1955). Analysis of dispersion for multiply classified data with unequal numbers in cells. *Sankhyā*, **15**, 253.

RAO, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrics*, **14**, 1.

ROY, S. N. & BOSE, R. C. (1953). Simultaneous confidence interval estimation. *Ann. Math. Statist.* **24**, 513.

SCHEFFÉ, H. (1953). A method for judging all contrasts in the analysis of variance. *Biometrika*, **40**, 87.

TUKEY, J. W. (1951). Quick and dirty methods in statistics Part II. Simple analysis for standard designs. *Proc. Fifth Annual Convention Amer. Soc. for Quality Control*, p. 189.

TESTS OF SIGNIFICANCE IN CANONICAL ANALYSIS

By D. N. LAWLEY

Mathematical Institute, University of Edinburgh

SUMMARY. Results are obtained concerning the approximate distribution of canonical correlation coefficients. A slight improvement is suggested for the test of significance of residual canonical roots.

1. Introduction

In this paper our object is to investigate in some detail the approximate distribution of canonical correlation coefficients and the tests of significance associated with them. In particular we shall discuss th. approximate test of significance proposed by Bartlett (1938, 1947) for the residual canonical roots when, let us say, the first k roots, corresponding to real relationships between two sets of variates, have been removed. Our approach involves some rather tedious and laborious algebra, but we have been unable to find any which is more elegant. No attempt has been made to justify the various expansions in series; they are presumed to be valid in an asymptotic sense.

2. Assumptions and notation

Let two sets of variates containing respectively p and q variates be denoted by column vectors x_1 and x_2, the corresponding row vectors being x_1' and x_2'. We shall be interested in the canonical relationships between x_1 and x_2, and we shall begin by considering the true correlation case where all variates follow a multivariate normal distribution with zero means. Without loss of generality, since at most only a linear transformation within each set is required, we may assume that their variances and covariances are given in matrix form by

$$\left.\begin{aligned} E(x_1x_1') = I_p, \quad E(x_2x_2') = I_q, \\ E(x_1x_2') = P, \quad E(x_2x_1') = P', \end{aligned}\right\} \tag{1}$$

where I_p denotes the unit matrix of order p and where $P = [\rho_{ij}]$ is a matrix with p rows and q columns such that $\rho_{ij} = 0$ if $i \neq j$ and such that $\rho_{ii} = \rho_i$ is the ith canonical correlation coefficient in order of magnitude. We shall assume that the first k ($< p$, $< q$) of the ρ_i are distinct and positive and that the remainder are zero.

Corresponding to the above true values we shall suppose that the usual sample estimates of the variances and covariances, with n degrees of freedom, are given by

$$\begin{aligned} A_{11} &= I_p + X_{11}, \quad A_{22} = I_q + X_{22}, \\ A_{12} &= P + X_{12}, \quad A_{21} = A_{12}' = P' + X_{21}. \end{aligned}$$

Thus the X-matrices represent deviations of sample estimates from expectation.

The squares $r_1^2, \ldots, r_q^2$, of the estimated canonical correlation coefficients are the latent roots of the matrix

$$A_{22}^{-1}A_{21}A_{11}^{-1}A_{12} = (I + X_{22})^{-1}(P' + X_{21})(I + X_{11})^{-1}(P + X_{12}). \tag{2}$$

If $p < q$, the last $q - p$ of these roots are identically zero. The usual difficulty arises with regard to the correspondence of these roots with their population values. We shall assume, however, that they are arranged in order of magnitude and that with high probability r_i^2 then

corresponds with ρ_i^2 for $i = 1, 2, \ldots, k$. It would be difficult to investigate the point with any degree of rigour, but it seems likely that, roughly speaking, we are justified in this assumption provided that the sample size is reasonably large and that the first k of the ρ_i are neither very near one another nor small.

Let us now expand the matrix product (2) in the form

$$(I - X_{22} + X_{22}^2 - \ldots)(P' + X_{21})(I - X_{11} + X_{11}^2 - \ldots)(P + X_{12}) = P'P + Y,$$

where $P'P$ is a diagonal matrix with elements $\rho_1^2, \ldots, \rho_q^2$. The matrix Y may be expressed as

$$Y = Y^{(1)} + Y^{(2)} + Y^{(3)} + \ldots,$$

where

$$\left.\begin{aligned} Y^{(1)} &= X_{22}(I - P'P) - X_{33}, \\ Y^{(2)} &= -X_{22}^2(I - P'P) + X_{22}X_{33} + X_{31}X_{13}, \\ Y^{(3)} &= X_{22}^3(I - P'P) - X_{22}^2X_{33} - X_{22}X_{31}X_{13} - X_{31}X_{11}X_{13}, \text{ etc.}, \end{aligned}\right\} \tag{3}$$

and where

$$\begin{aligned} X_{33} &= X_{22} - P'X_{12} - X_{21}P + P'X_{11}P, \\ X_{31} &= X_{13}' = X_{21} - P'X_{11}. \end{aligned}$$

In order to interpret the above notation let us introduce a vector x_3 of variates given by

$$x_3 = x_2 - P'x_1.$$

Thus x_3 represents the component of x_2 which is independent of x_1. The variances and covariances involving x_3 are then given by

$$\left.\begin{aligned} E(x_3x_3') &= E(x_3x_2') = I - P'P, \\ E(x_3x_1') &= 0 \end{aligned}\right\} \tag{4}$$

The matrices X_{33} and X_{31} are to be interpreted as representing corresponding deviations of sample estimates from expectation.

3. Various expectations

We now require the expectations of elements and various products of elements of the matrices $Y^{(1)}$, $Y^{(2)}$, $Y^{(3)}$ and of Y. By the use of equations (3) we can express these in terms of the expectations of products of elements of the X-matrices. The latter can be evaluated by the usual combinatorial rules and by the use of the relations (1) and (4). The algebra is, however, extremely heavy, so we shall not reproduce it in full. Let us denote the typical elements of Y and of $Y^{(r)}$ by y_{ij} and $y_{ij}^{(r)}$ respectively $(i, j = 1, 2 \ldots, q)$. Then

$$y_{ij} = y_{ij}^{(1)} + y_{ij}^{(2)} + y_{ij}^{(3)} + \ldots,$$

and we find, first, that

$$E(y_{ii}) = (1 - \rho_i^2)(p - a_1 - \rho_i^2)/n + O(n^{-2}),$$

where

$$a_1 = \sum_{j=1}^{k} \rho_j^2.$$

For $i \neq j$ we have $E(y_{ij}) = 0$.

Proceeding further, for $i \neq j$, we have

$$\begin{aligned} E(y_{ij}^{(1)}y_{ji}^{(1)}) &= (1 - \rho_i^2)(1 - \rho_j^2)(\rho_i^2 + \rho_j^2)/n, \\ E(y_{ij}^{(1)}y_{ji}^{(2)}) &= -(1 - \rho_i^2)(1 - \rho_j^2)\{(q + 2)\rho_i^2 + p\rho_j^2\}/n^2, \\ E(y_{ij}^{(1)}y_{ji}^{(3)}) &= (1 - \rho_i^2)(1 - \rho_j^2)\{(q - p + a_1 + 1)\rho_i^2 + (a_1 - 1)\rho_j^2 + \rho_i^4 + \rho_j^4 + 2\rho_i^2\rho_j^2\}/n^2, \\ E(y_{ij}^{(2)}y_{ji}^{(2)}) &= (1 - \rho_i^2)(1 - \rho_j^2)\{(p - 2a_1 + a_2) + (a_1 - 2)(\rho_i^2 + \rho_j^2) + 2(\rho_i^4 + \rho_j^4 + \rho_i^2\rho_j^2)\}/n^2, \end{aligned}$$

where

$$a_2 = \sum_{s=1}^{k} \rho_s^4.$$

We hence find that

$$E(y_{ij}y_{ji})/(1-\rho_i^2)(1-\rho_j^2) = (\rho_i^2+\rho_j^2)/n + \{(p-2a_1+a_2)+(3a_1-2p+4)(\rho_i^2+\rho_j^2) + 2(2\rho_i^4+2\rho_j^4+3\rho_i^2\rho_j^2)\}/n^2+O(n^{-3}).$$

In a similar manner we obtain the following results:

$$\text{var}(y_{ii}) = 4\rho_i^2(1-\rho_i^2)^2/n+(1-\rho_i^2)(q-a_1)/n^2+(1-\rho_i^2)^2\{(2p-q-2a_1+a_2) + 2(6a_1-4p-11)\rho_i^2+39\rho_i^4\}/n^2+O(n^{-3}),$$

$$\kappa_3(y_{ii}) = 24\rho_i^2(1-\rho_i^2)^3(1-3\rho_i^2)/n^2+O(n^{-3}),$$

$$\kappa_4(y_{ii}) = 96\rho_i^2(1-\rho_i^2)^4(1-2\rho_i^2)(2-13\rho_i^2)/n^3+O(n^{-4}).$$

Also, for $i \neq j$, we have

$$\text{var}(y_{ij}) = (1-\rho_j^2)(\rho_i^2+\rho_j^2-2\rho_i^2\rho_j^2)/n+O(n^{-2}),$$

$$\text{cov}(y_{ii},y_{jj}) = (1-\rho_i^2)(1-\rho_j^2)\{-5(\rho_i^2+\rho_j^2)+4(\rho_i^4+\rho_j^4)+5\rho_i^2\rho_j^2\}/n^2+O(n^{-3}),$$

$$\text{cov}\{y_{ii},(y_{ij}y_{ji})\} = (1-\rho_i^2)^2(1-\rho_j^2)(1+5\rho_i^2+\rho_j^2-8\rho_i^4-\rho_j^4-9\rho_i^2\rho_j^2)/n^2-(1-\rho_i^2)(1-\rho_j^2)^2/n^2+O(n^{-3}).$$

4. Approximate distribution of the first k roots

We now make use of an expression for the first k latent roots of the matrix $P'P+Y$. A similar expansion was derived in a previous paper (Lawley, 1956*a*) in which we were concerned with the internal analysis of a single set of variates. The sth root of the matrix ($s = 1, 2, \ldots, k$) is given by

$$r_s^2 = \rho_s^2+y_{ss}+\sum_i{}' \frac{y_{si}y_{is}}{\rho_s^2-\rho_i^2}-y_{ss}\sum_i{}' \frac{y_{si}y_{is}}{(\rho_s^2-\rho_i^2)^2}+\sum_{i,j}{}' \frac{y_{si}y_{ij}y_{js}}{(\rho_s^2-\rho_i^2)(\rho_s^2-\rho_j^2)}+O(y^4),$$

where Σ' indicates summation over all values from 1 to q with the exception of s. The validity of the expansion depends upon the assumption that the y-variates, which are of order $1/\sqrt{n}$, are small compared with the quantities $\rho_s^2-\rho_i^2$.

By means of this expansion and of the results of the preceding section we are able to evaluate the first four cumulants of r_s^2 and hence those of r_s. We find that

$$2\rho_s E(r_s-\rho_s) = (1-\rho_s^2)\left\{p+q-2-\rho_s^2+2(1-\rho_s^2)\sum_i{}'\left(\frac{\rho_i^2}{\rho_s^2-\rho_i^2}\right)\right\}\Big/n+O(n^{-2}),$$

$$\text{var}(r_s) = (1-\rho_s^2)^2/n+O(n^{-2}).$$

$$\kappa_3(r_s) = -6\rho_s(1-\rho_s^2)^3/n^2+O(n^{-3}),$$

$$\kappa_4(r_s) = -6(1-\rho_s^2)^4(1-12\rho_s^2)/n^3+O(n^{-4}).$$

In general, it seems useless to seek a better approximation to the variance of r_s, since the term of order n^{-2} is extremely cumbersome. The variance depends to a considerable extent not only on the value of ρ_s but also on the differences $\rho_s^2-\rho_i^2$. It is, however, worth considering the case where $k = 1$, i.e. where ρ_1 is the only non-zero canonical correlation coefficient. More generally, for $k > 1$, we can consider r_k under the somewhat restrictive assumption that the first $k-1$ of the ρ_i are near unity while ρ_k is moderate. For these cases, writing ρ in place of ρ_k we have

$$E(r_k) \approx \rho+\tfrac{1}{2}(1-\rho^2)(m\rho^{-1}-\rho)/n,$$

$$\text{var}(r_k) \approx (1-\rho^2)^2/n+\tfrac{1}{2}(1-\rho^2)^2\{11\rho^2-2(m-k+1)-m\rho^{-2}\}/n^2,$$

where $m = p+q-2k$.

It is natural to apply the usual z-transformation in order to normalize the distributions of the canonical correlation coefficients; but in general this is unfortunately of little help

as it fails to stabilize the mean and variance to any marked extent. The transformation is only worth while if we are prepared to make the same assumptions as in the preceding paragraph, in which case, writing r for r_k, let us put

$$z = \tfrac{1}{2}\log_e \frac{1+r}{1-r}, \quad \zeta = \tfrac{1}{2}\log_e \frac{1+\rho}{1-\rho}.$$

Then since

$$z-\zeta = x+\rho x^2+\tfrac{1}{3}(1+3\rho^2)\,x^3+\ldots, \quad \text{where} \quad x = (r-\rho)/(1-\rho^2),$$

it is easy to show that

$$E(z) \approx \zeta+\tfrac{1}{2}(\rho+m\rho^{-1})/n,$$
$$\kappa_2(z) \approx 1/(n-p-q+k-1+\tfrac{1}{2}\rho^2+\tfrac{1}{2}m\rho^{-2}),$$
$$\kappa_3(z) = O(n^{-3}), \quad \kappa_4(z) = 2/n^3+O(n^{-4}).$$

It is of particular interest to consider the multiple correlation coefficient between one variate and another set of p variates. In this case, putting $q = 1$, $k = 1$ and $m = p-1$, we have

$$E(z) \approx \zeta+\tfrac{1}{2}\{\rho+(p-1)\,\rho^{-1}\}/n,$$
$$\text{var}\,(z) \approx 1/\{n-p-1+\tfrac{1}{2}\rho^2+\tfrac{1}{2}(p-1)\,\rho^{-2}\}.$$

The mean and variance of z are not as well stabilized as in the case of an ordinary correlation coefficient, and the results clearly break down if ρ tends to zero. For practical use it would be necessary to substitute r for ρ in these formulae.

Though the result has little practical importance, it may be shown that the correlation coefficient between r_i and r_j $(i < j \leqslant k)$ is

$$\frac{2\rho_i\rho_j(1-\rho_i^2)\,(1-\rho_j^2)}{n(\rho_i^2-\rho_j^2)^2}+O(n^{-2}).$$

This breaks down if either ρ_i^2, ρ_j^2 or the difference between them is small.

5. Test of significance of the residual roots

The residual roots $r_{k+1}^2, \ldots, r_q^2$ are the $q-k$ smallest roots of the matrix $P'P+Y$. Let us partition this matrix in such a way as to divide the first k rows and columns from the remaining $q-k$ rows and columns. Then we may write it as

$$\begin{bmatrix} D+Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix},$$

where D is a diagonal matrix with elements $\rho_1^2, \ldots, \rho_k^2$.

By following the method employed in our previous (1956*a*) paper we can construct a matrix of $q-k$ rows and columns whose latent roots differ from $r_{k+1}^2, \ldots, r_q^2$ only by terms of the fifth or higher degree in the elements of Y. This matrix is given by

$$\begin{aligned} Z = {} & Y_{22}-Y_{21}D^{-1}Y_{12}-Y_{22}Y_{21}D^{-2}Y_{12}+Y_{21}D^{-1}Y_{11}D^{-1}Y_{12}-Y_{22}Y_{21}D^{-3}Y_{12}Y_{22} \\ & +Y_{22}Y_{21}D^{-2}Y_{11}D^{-1}Y_{12}+Y_{22}Y_{21}D^{-1}Y_{11}D^{-2}Y_{12}-Y_{21}D^{-1}Y_{11}D^{-1}Y_{11}D^{-1}Y_{12} \\ & +Y_{21}D^{-1}Y_{12}Y_{21}D^{-2}Y_{12}. \end{aligned} \tag{5}$$

For our present purpose it is necessary to carry out the expansion as far as terms of the fourth degree, whereas the corresponding expansion used formerly was taken only as far as terms of the third degree. The validity of the expansion depends on the fact that the elements of Y are small in probability compared with ρ_k^2.

The log-likelihood ratio criterion used by Bartlett for testing the hypothesis that $\rho_{k+1},\ldots,\rho_q$ are all zero is, apart from a multiplying factor,

$$-\log_e \prod_{i=k+1}^{q} (1-r_i^2). \tag{6}$$

The product is that of the latent roots of the matrix $I-Z$, i.e. is the determinant of $I-Z$, to the order of approximation stated above. Hence the criterion may be written as

$$-\log_e |I-Z| = \operatorname{tr}(Z) + \tfrac{1}{2}\operatorname{tr}(Z^2) + \tfrac{1}{3}\operatorname{tr}(Z^3) + \ldots, \tag{7}$$

where tr denotes trace or sum of diagonal elements. This, when multiplied by a suitable factor, is distributed approximately as χ^2 with $(p-k)(q-k)$ degrees of freedom. The factor will be determined by finding the expectation of (7) correct to terms of order n^{-2}.

For this purpose we require the results given below. For the first of these the matrix Y requires expansion as far as the component $Y^{(4)}$. Assuming throughout that $\rho_s = 0 (k < s \leqslant q)$ and neglecting quantities of order n^{-3} we have

$$\begin{aligned} E(y_{ss}) &= (p-a_1)/n + \{(p+1)a_1 - a_1^2 - a_2\}/n^2, \\ E(y_{ss}^2) &= \{(2p-3a_1+a_2) + (p-a_1)^2\}/n^2. \end{aligned}$$

For $i \neq s$ we have

$$\begin{aligned} E(y_{si}y_{is}) &= \rho_i^2(1-\rho_i^2)/n + (1-\rho_i^2)\{(p-2a_1+a_2) - (2p-3a_1+4)\rho_i^2 + 4\rho_i^4\}/n^2, \\ E(y_{ss}y_{si}y_{is}) &= \rho_i^2(1-\rho_i^2)(p-a_1+2-\rho_i^2)/n^2, \\ E(y_{si}y_{ii}y_{is}) &= \rho_i^2(1-\rho_i^2)\{(p-a_1+4)(1-\rho_i^2) - 10\rho_i^2 + 9\rho_i^4\}/n^2, \\ E(y_{si}^2y_{is}^2) &= \rho_i^4(1-\rho_i^2)(3-2\rho_i^2)/n^2, \\ E(y_{si}y_{ii}^2y_{is}) &= 4\rho_i^4(1-\rho_i^2)^3/n^2; \end{aligned}$$

and, for i, j, s all unequal,

$$\begin{aligned} E(y_{si}y_{ij}y_{js}) &= (1-\rho_i^2)(1-\rho_j^2)(\rho_i^2+\rho_j^2)(1-\rho_i^2-\rho_j^2)/n^2, \\ E(y_{si}y_{ij}y_{ji}y_{is}) &= \rho_i^2(1-\rho_i^2)(1-\rho_j^2)(\rho_i^2+\rho_j^2)/n^2, \\ E(y_{si}y_{is}y_{sj}y_{js}) &= \rho_i^2\rho_j^2(1-\rho_i^2)(1-\rho_j^2)/n^2. \end{aligned}$$

By the use of these results we find, to the same order of approximation, that

$$\begin{aligned} &E(\operatorname{tr} Y_{22}) = (q-k)(p-a_1)/n + (q-k)\{(p+1)a_1 - a_1^2 - a_2\}/n^2, \\ &E(\operatorname{tr} Y_{21}D^{-1}Y_{12}) = (q-k)(k-a_1)/n + (q-k)\{(b-k)(p-2a_1+a_2) - (k-a_1)(2p-3a_1) \\ &\qquad\qquad - 4(k-2a_1+a_2)\}/n^2, \\ &E(\operatorname{tr} Y_{22}Y_{21}D^{-2}Y_{12}) = (q-k)\{(q-k)(b-2k+a_1) + (b-k)(p-a_1+1)\}/n^2, \\ &E(\operatorname{tr} Y_{21}D^{-1}Y_{11}D^{-1}Y_{12}) = (q-k)\{2(b-k)(k-2a_1+a_2) - 2(k-a_1)^2 + (b-2k+a_1)(p-a_1+2) \\ &\qquad\qquad - 6k + 11a_1 - 5a_2\}/n^2, \\ &E(\operatorname{tr} Y_{21}D^{-1}Y_{11}D^{-1}Y_{11}D^{-1}Y_{12}) = (q-k)\{(k-a_1)(b-2k+a_1) + (b-k)(k-2a_1+a_2) \\ &\qquad\qquad + 2(b-3k+3a_1-a_2)\}/n^2, \\ &E(\operatorname{tr} Y_{21}D^{-1}Y_{12}\,Y_{21}D^{-2}Y_{12}) = (q-k)\{(q-k)(b-2k+a_1) + (b-k)(k-a_1+1)\}/n^2, \end{aligned}$$

where

$$b = \sum_{j=1}^{k} (1/\rho_j^2).$$

The expectations of the remaining terms in tr Z may be neglected, and hence

$$E(\operatorname{tr} Z) = (p-k)(q-k)\{1/n + (2k-b)/n^2\}.$$

The only part of tr Z^2 with which we need be concerned is

$$\operatorname{tr}\{Y_{22}^2 - 2Y_{22}Y_{21}D^{-1}Y_{12} + (Y_{21}D^{-1}Y_{12})^2\},$$

since the expectations of all other terms are of order less than n^{-2} and may be neglected. We find that

$$E(\operatorname{tr} Y_{22}^2) = (q-k)\{(q-k)(p-2a_1+a_2)+(p-a_1)(p-a_1+1)\}/n^2,$$
$$E(\operatorname{tr} Y_{22}Y_{21}D^{-1}Y_{12}) = (q-k)\{(q-k)(k-2a_1+a_2)+(k-a_1)(p-a_1+1)\}/n^2,$$
$$E\{\operatorname{tr}(Y_{21}D^{-1}Y_{12})^2\} = (q-k)\{(q-k)(k-2a_1+a_2)+(k-a_1)(k-a_1+1)\}/n^2,$$

and hence that

$$E(\operatorname{tr} Z^2) = (p-k)(q-k)(p+q-2k+1)/n^2.$$

The expectations of $\operatorname{tr} Z^3$, $\operatorname{tr} Z^4$, etc., are all of order less than n^{-2} and may be neglected. Hence, finally the expectation of (7) is found to be

$$(p-k)(q-k)\{1/n+\tfrac{1}{2}(p+q+2k-2b+1)/n^2\}+O(n^{-3}).$$

Thus, supposing that the values of $\rho_1, \ldots, \rho_k$ were known, the correct multiplying factor for the criterion (6) would be

$$n-k-\tfrac{1}{2}(p+q+1)+\sum_{j=1}^{k}(1/\rho_j^2).$$

This would not only make the expectation of the criterion equal to $(p-k)(q-k)+O(n^{-2})$, but would make the other moments equal to those of χ^2 apart from errors of order n^{-2}, by a result established elsewhere (Lawley, 1956*b*).

If $\rho_1, \ldots, \rho_k$ are all fairly near unity, the multiplying factor is approximately $n-\frac{1}{2}(p+q+1)$, which is the value suggested by Bartlett. Otherwise an improvement should, presumably, be obtained by taking the factor as

$$n-k-\tfrac{1}{2}(p+q+1)+\sum_{j=1}^{k}(1/r_j^2), \tag{8}$$

though we cannot claim this to be very firmly established.

6. The case of fixed independent variates

Let us now suppose that x_1 represents a set of fixed independent variates and x_2 the dependent variates. Then the canonical analysis is to be regarded as a regression analysis of x_2 on x_1. For given x_1 the variates x_2 are assumed to be normally distributed with constant variances and covariances and with mean values given by the vector $B'x_1$, where $B = [\beta_{ij}]$ is a $p \times q$ matrix whose elements represent coefficients of linear regression. Without loss of generality we may assume that the variates are chosen in such a way that $\beta_{ij} = 0$ if $i \neq j$ and that $\beta_{ii} = \beta_i$, where $\beta_1, \ldots, \beta_k$ are positive and arranged in order of magnitude while the remaining β_i are zero. We may also assume that, for given x_1, the variates x_2 have unit variances and zero covariances. These assumptions require a change of notation.

In considering sample values it is now convenient to use sums of squares and products (with n degrees of freedom). These form matrices which we denote by $C_{11}, C_{12} = C_{21}'$ and C_{22}. It will be convenient to consider the matrix

$$(C_{22}-C_{21}C_{11}^{-1}C_{12})^{-1}C_{21}C_{11}^{-1}C_{12}, \tag{9}$$

which has for its latent roots the quantities $r_i^2/(1-r_i^2)$, where $i = 1, 2, \ldots, q$.

Estimates, with $m = n-p$ degrees of freedom, of the variances and covariances of the x_2 (for given x_1) are provided by the elements of the matrix

$$\frac{1}{m}(C_{22}-C_{21}C_{11}^{-1}C_{12}) = I_q + V,$$

where the elements of V represent deviations of these estimates from expectation. To simplify matters we may choose the x_1, again without loss of generality, in such a way that $C_{11} = mI_p$. Then we may write

$$\frac{1}{m}C_{12} = C_{11}^{-1}C_{12} = B + X,$$

where the elements of the $p \times q$ matrix X represent deviations of estimated regression coefficients from expectation. They constitute a set of pq independent normal variates each having zero mean and variance $1/m$ and they are distributed independently of the elements of V.

The matrix given by (9) may now be expressed in the form

$$(I+V)^{-1}(B'+X')(B+X) = (I-V+V^2-\ldots)(B'+X')(B+X) = B'B+Y,$$

where $B'B$ is a diagonal matrix with elements $\beta_i^2 (i = 1, 2, \ldots, q)$. The first k roots of (9) provide estimates of $\beta_1^2, \ldots, \beta_k^2$, while the remainder enable us to test the hypothesis that $\beta_{k+1}, \ldots, \beta_q$ are all zero. The matrix Y may be expanded as before in the form

$$Y = Y^{(1)} + Y^{(2)} + Y^{(3)} + \ldots,$$

but now

$$\begin{aligned} Y^{(1)} &= B'X + X'B - VB'B, \\ Y^{(2)} &= X'X - VB'X - VX'B + V^2B'B, \\ Y^{(3)} &= -VY^{(2)}, \\ Y^{(4)} &= V^2Y^{(2)}, \text{ etc.} \end{aligned}$$

Let us partition Y in the same manner as the Y of the preceding section. Then the $q-k$ smallest roots of $B'B+Y$ are approximately the same as those of the matrix Z, where Z is as given by (5), provided that we now define D as the diagonal matrix with elements $\beta_1^2, \ldots, \beta_k^2$. The latent roots of $I+Z$ are approximately equal to the quantities $1/(1-r_i^2)$, for $i = k+1, \ldots, q$. Hence the criterion (6) may in this case be expressed as

$$\log_e |I+Z| = \operatorname{tr}(Z) - \tfrac{1}{2}\operatorname{tr}(Z^2) + \tfrac{1}{3}\operatorname{tr}(Z^3) - \ldots. \qquad (10)$$

In order to evaluate the expectation of this expression it is necessary to recalculate the expectations of the various products of elements of Y. This is however a comparatively simple process. Assuming throughout that $\beta_s = 0 (k < s \leqslant q)$ and neglecting quantities of order m^{-3} we find that

$$\begin{aligned} E(y_{ss}) &= p/m + p(q+1)/m^2, \\ E(y_{ss}^2) &= (p^2+2p+c)/m^2, \end{aligned}$$

where

$$c = \sum_{j=1}^{k} \beta_j^2.$$

For $i \neq s$ we have

$$\begin{aligned} E(y_{si}y_{is}) &= \beta_i^2/m + \{p + (p+2q+2)\beta_i^2\}/m^2, \\ E(y_{ss}y_{si}y_{is}) &= \beta_i^2(p+2+\beta_i^2)/m^2, \\ E(y_{si}y_{ii}y_{is}) &= \beta_i^2\{(p+4)+(q+4)\beta_i^2\}/m^2, \\ E(y_{si}^2y_{is}^2) &= \beta_i^4(3+\beta_i^2)/m^2, \\ E(y_{si}y_{ii}^2y_{is}) &= 2\beta_i^4(2+\beta_i^2)/m^2; \end{aligned}$$

and, for i, j, s all unequal,

$$\begin{aligned} E(y_{si}y_{ij}y_{js}) &= (\beta_i^2+\beta_j^2+2\beta_i^2\beta_j^2)/m^2, \\ E(y_{si}y_{ij}y_{ji}y_{is}) &= \beta_i^2(\beta_i^2+\beta_j^2+\beta_i^2\beta_j^2)/m^2, \\ E(y_{si}y_{is}y_{sj}y_{js}) &= \beta_i^2\beta_j^2/m^2. \end{aligned}$$

We then have

$$\begin{aligned}
E(\operatorname{tr} Y_{22}) &= p(q-k)/m + p(q-k)(q+1)/m^2,\\
E(\operatorname{tr} Y_{21}D^{-1}Y_{12}) &= k(q-k)/m + (q-k)\{pd + k(p+2q+2)\}/m^2,\\
E(\operatorname{tr} Y_{22}Y_{21}D^{-2}Y_{12}) &= (q-k)\{(p+q-k+1)d + k\}/m^2,\\
E(\operatorname{tr} Y_{21}D^{-1}Y_{11}D^{-1}Y_{12}) &= (q-k)\{(p+2k+2)d + k(q+2k+2)\}/m^2,\\
E(\operatorname{tr} Y_{21}D^{-1}Y_{11}D^{-1}Y_{11}D^{-1}Y_{12}) &= (q-k)(k+1)(k+2d)/m^2,\\
E(\operatorname{tr} Y_{21}D^{-1}Y_{12}Y_{21}D^{-2}Y_{12}) &= (q-k)\{(q+1)d + k\}/m^2,
\end{aligned}$$

where

$$d = \sum_{j=1}^{k}(1/\beta_j^2).$$

Hence

$$E(\operatorname{tr} Z) = (p-k)(q-k)\{1/m + (q-k+1-d)/m^2\}.$$

We have also

$$\begin{aligned}
E(\operatorname{tr} Y_{22}^2) &= (q-k)\{p(p+q-k+1)+c\}/m^2,\\
E(\operatorname{tr} Y_{22}Y_{21}D^{-1}Y_{12}) &= (q-k)\{k(p+q-k+1)+c\}/m^2,\\
E\{\operatorname{tr}(Y_{21}D^{-1}Y_{12})^2\} &= (q-k)\{k(q+1)+c\}/m^2;
\end{aligned}$$

and thus

$$E(\operatorname{tr} Z^2) = (p-k)(q-k)(p+q-2k+1)/m^2.$$

Hence, finally, the expectation of (10) is found to be

$$\begin{aligned}
&(p-k)(q-k)\{1/m + \tfrac{1}{2}(q-p-2d+1)/m^2\} + O(m^{-3})\\
&\qquad = (p-k)(q-k)\{1/n + \tfrac{1}{2}(p+q-2d+1)/n^2\} + O(n^{-3}),
\end{aligned}$$

putting $m = n-p$.

The appropriate multiplying factor for the criterion (6) is

$$n - \tfrac{1}{2}(p+q+1) + \sum_{j=1}^{k}(1/\beta_j^2).$$

If $\beta_1, \ldots, \beta_k$ are all large the last term may be neglected. If, on the other hand, we replace each β_j^2 by its estimate $r_j^2/(1-r_j^2)$, the multiplying factor will again be given by expression (8). We are thus led to much the same conclusion as in the true correlation case.

REFERENCES

Bartlett, M. S. (1938). Further aspects of the theory of multiple regression. *Proc. Camb. Phil. Soc.* **34**, 33–40.

Bartlett, M. S. (1947). Multivariate analysis. *J.R. Statist. Soc.* (Suppl.), **9**, 176–90.

Lawley, D. N. (1956*a*). Tests of significance for the latent roots of covariance and correlation matrices. *Biometrika*, **43**, 128–36.

Lawley, D. N. (1956*b*). A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika*, **43**, 295–303.

THE ESTIMATION OF PARAMETERS IN SYSTEMS OF STOCHASTIC DIFFERENTIAL EQUATIONS

By A. W. PHILLIPS

London School of Economics

1. Summary

In many fields of study there occur systems of lagged dependences which can be described approximately by systems of linear stationary stochastic differential equations. In this paper we consider the estimation of parameters in such systems, using an approach similar to that of Bartlett (1946, 1955) and Quenouille (1957), and discuss conditions for identification of the separate behaviour relationships of a system. Little is known about the small-sample properties of estimates in systems of this kind, but sampling experiments are being carried out and the results will be published later.

2. Systems of linear stochastic lagged dependences

A system of linear stochastic lagged dependences between n variables $x_i(t)$ which are functions of continuous time can be described by the system of convolution integrals

$$x_i(t) = \sum_{j \neq i} \int_0^\infty w_{ij}(h)\, x_j(t-h)\, dh + \xi_i(t) \quad (i,j = 1, 2, \ldots, n), \tag{1}$$

in which each of the weighting functions $w_{ij}(h)$ specifies the magnitude of the dependence of the value taken by x_i at time t on the values taken by x_j at times $t-h$, and the $\xi_i(t)$ are stochastic disturbance terms. The Laplace transform of the weighting function of a lagged dependence is called the transfer function of the dependence. If all the transfer functions of the system (1) are rational functions, i.e. if

$$\int_0^\infty e^{-sh} w_{ij}(h)\, dh = \frac{\beta_{ij}(s)}{\alpha_{ij}(s)}, \tag{2}$$

where $\beta_{ij}(s)$ and $\alpha_{ij}(s)$ are polynomials of finite degree in s, (1) can be written as the system of differential equations

$$x_i(t) = \sum_{j \neq i} \frac{\beta_{ij}(D)}{\alpha_{ij}(D)} x_j(t) + \xi_i(t) \quad (i,j = 1, 2, \ldots, n) \tag{3}$$

in which D is the differential operator d/dt. For the ith equation in (1) is then a particular integral of the ith equation in (3), and if the system is stable and the initial conditions are at $t = -\infty$ it is the general solution of the ith equation in (3).

Let $Q(D)$ be an $n \times n$ matrix whose elements $Q_{ij}(D)$ are given by

$$Q_{ij}(D) = \begin{cases} -\dfrac{\beta_{ij}(D)}{\alpha_{ij}(D)} & \text{if } j \neq i, \\ 1 & \text{if } j = i \end{cases} \tag{4}$$

and let $x(t)$ and $\xi(t)$ be column vectors with elements $x_i(t)$ and $\xi_i(t)$. Then (3) may be written

$$Q(D)\, x(t) = \xi(t). \tag{5}$$

In many fields of study a system like (3) may be used to represent the behaviour of a number of separate units which form an interdependent system. Thus the equations may represent the behaviour of units of physical equipment, as in control engineering, or of groups of persons or firms, as in economics, or of animals or organisms, as in ecological or biological studies. In describing real behaviour systems of this sort it is often reasonable to assume that the behaviour units cannot respond to change instantaneously. This is equivalent to the assumption that the weighting functions $w_{ij}(h)$ in (1) do not contain delta-function impulse terms,* which in turn implies that the numerator polynomial $\beta_{ij}(s)$ of each transfer function is of lower degree than the denominator polynomial $\alpha_{ij}(s)$. In this paper we shall consider only systems in which the numerator of each transfer function is of lower degree than the denominator. It will be seen later that this restriction plays an important part in the identification of the separate behaviour equations of a system.

3. Specification of the stochastic disturbances $\xi(t)$

We shall consider only systems which are stable and in which the disturbances $\xi(t)$ are stationary stochastic variables whose covariance functions possess rational two-sided Laplace transforms. It will be convenient to assume that the disturbances $\xi(t)$ are generated from another set of stationary stochastic variables $\zeta_k(t)$ by the transformation

$$\xi_i(t) = \sum_{k=1}^{n} \int_0^\infty r_{ik}(h)\,\zeta_k(t-h)\,dh \quad (i = 1, 2, \ldots, n) \tag{6}$$

in which the functions $r_{ik}(h)$ have Laplace transforms $R_{ik}(s)$ of the form

$$R_{ik}(s) = \frac{\eta_{ik}(s)}{\gamma_{ik}(s)}, \tag{7}$$

where $\eta_{ik}(s)$ and $\gamma_{ik}(s)$ are polynomials in s, the degree of $\eta_{ik}(s)$ being lower than that of $\gamma_{ik}(s)$. Equations (6) can then be written as the set of differential equations

$$\xi(t) = R(D)\,\zeta(t), \tag{8}$$

where $R(D)$ is a matrix with elements $R_{ik}(D)$ and $\zeta(t)$ is a column vector with elements $\zeta_k(t)$. In order that the variables $\xi(t)$ be stationary it is necessary that all the roots of the equations $\gamma_{ik}(s) = 0$ have negative real parts.

We define the following covariance functions;

$$\left.\begin{aligned} \vartheta(\tau) &= E\{\xi(t)\,\xi'(t-\tau)\}, \\ \psi(\tau) &= E\{\zeta(t)\,\zeta'(t-\tau)\}, \\ \lambda(\tau) &= E\{\zeta(t)\,\xi'(t-\tau)\} \end{aligned}\right\} \tag{9}$$

* The δ-function or unit impulse function, $\delta(h)$ is defined by the equations

$$\delta(h) = \begin{cases} 0 & \text{if } h \neq 0 \\ \infty & \text{if } h = 0 \end{cases} \quad \text{and} \quad \int_{-\infty}^{\infty} \delta(h)\,dh = 1.$$

(See James, Nichols & Phillips (1947), chapter 2, especially § 2·4.) The Laplace transform of $\delta(h)$ is clearly unity. If $\beta_{ij}(s)$ were not of lower degree than $\alpha_{ij}(s)$ the partial fraction expansion of $\beta_{ij}(s)/\alpha_{ij}(s)$ would contain a constant term and so its inverse transform $w_{ij}(h)$ would contain a δ-function impulse term.

and denote their Laplace transforms by $\theta(s)$, $\Psi(s)$ and $\Lambda(s)$. Thus

$$\left.\begin{aligned}\theta(s) &= \int_{-\infty}^{\infty} e^{-s\tau}\vartheta(\tau)\,d\tau,\\ \Psi(s) &= \int_{-\infty}^{\infty} e^{-s\tau}\psi(\tau)\,d\tau,\\ \Lambda(s) &= \int_{-\infty}^{\infty} e^{-s\tau}\lambda(\tau)\,d\tau.\end{aligned}\right\} \tag{10}$$

It will now be shown that $\theta(s)$ is given by

$$\theta(s) = R(s)\Psi(s)\,R'(-s). \tag{11}$$

The functions $r_{ik}(h)$ in (6) may be regarded as defined but zero for $h < 0$, and the lower limit of integration extended to $-\infty$. In matrix notation (6) then becomes

$$\xi(t) = \int_{-\infty}^{\infty} r(h)\,\zeta(t-h)\,dh, \tag{12}$$

$r(h)$ being the matrix with elements $r_{ik}(h)$. Post-multiplying (12) by $\xi'(t-\tau)$ and taking expectations we obtain using (9),

$$\vartheta(\tau) = \int_{-\infty}^{\infty} r(h)\,\lambda(\tau-h)\,dh. \tag{13}$$

Taking the Laplace transform of (13) and remembering that the transform of the convolution of two functions is the product of their transforms* we have

$$\theta(s) = R(s)\,\Lambda(s). \tag{14}$$

Also, postmultiplying (12) by $\zeta'(t+\tau)$ and taking expectations we find, using (9)

$$\lambda'(\tau) = \int_{-\infty}^{\infty} r(h)\,\psi'(\tau+h)\,dh. \tag{15}$$

On making the substitution $h_1 = -h$ this becomes

$$\lambda'(\tau) = \int_{-\infty}^{\infty} r(-h_1)\,\psi'(\tau-h_1)\,dh_1 \tag{16}$$

and it follows that

$$\Lambda'(s) = R(-s)\,\Psi'(s). \tag{17}$$

Substituting the transpose of (17) in (14) we obtain (11).

We now postulate that $\Psi(s)$ be identically the unit matrix,

$$\Psi(s) = I. \tag{18}$$

Substitution of (18) in (11) then gives

$$\theta(s) = R(s)\,R'(-s). \tag{19}$$

Since the spectral density and cross-spectral density functions of a set of variables can be obtained from the two-sided Laplace transforms of their covariance functions by replacing

* See Widder (1946), p. 258, theorem 16*b*. The conditions for the theorem are satisfied as a result of the stationarity of $\zeta(t)$ and the conditions that the roots of the equations $\gamma_{ik}(s) = 0$ have negative real parts.

the complex number s by the pure imaginary $i\omega$ (the Laplace transform then reducing to a Fourier transform), the spectral densities of the variables $\zeta(t)$ are constants for all frequencies ω and their cross-spectral densities are identically zero. The matrix $\psi(\tau)$ of their covariance functions is diagonal, the leading elements being δ-function impulses. Since this implies that the variances are infinite, the variables $\zeta(t)$ are not physically realizable.* However, (6) remains a valid mathematical model for the generation of the physically realizable variables $\xi(t)$, whose spectral density and cross-spectral density functions, given by $\theta(i\omega)$, are proper rational functions of $i\omega$ as a result of the restriction that $\eta_{ik}(s)$ in (7) be of lower degree than $\gamma_{ik}(s)$.

4. Covariance functions of the system variables $x(t)$

Substituting (8) in (5) we have

$$Q(D)\,x(t) = R(D)\,\zeta(t). \tag{20}$$

Let $P(D)$ be a diagonal matrix whose ith leading element is the lowest common multiple of the denominators of the elements in the ith row of $Q(D)$ and in the ith row of $R(D)$, i.e. the lowest common multiple of the polynomials $\alpha_{ij}(D)$, $j \neq i$, in (4) and $\gamma_{ik}(D)$, $k = 1, 2, \ldots, n$ in (7), and let $F(D)$ and $G(D)$ be matrices defined by

$$F(D) = P(D)\,Q(D) \tag{21}$$

and

$$G(D) = P(D)\,R(D). \tag{22}$$

Then premultiplying (20) by $P(D)$ we obtain

$$F(D)\,x(t) = G(D)\,\zeta(t). \tag{23}$$

The elements of $F(D)$ and $G(D)$ are polynomials in D, the diagonal element in any row of $F(D)$ being of higher degree than the non-diagonal elements in the same row and also of higher degree than the elements in the same row of $G(D)$. Thus the determinant $|F(D)|$ cannot vanish identically, and we may write (23) as

$$x(t) = F^{-1}(D)\,G(D)\,\zeta(t). \tag{24}$$

Let $\phi(\tau)$ be the matrix of covariance functions of the system variables $x(t)$, so that

$$\phi(\tau) = E\{x(t)\,x'(t-\tau)\} \tag{25}$$

and let $\Phi(s)$ be the Laplace transform of $\phi(\tau)$. Then, by a similar argument to that used in § 3, we have

$$\Phi(s) = F^{-1}(s)\,G(s)\Psi(s)\,G'(-s)\,F'^{-1}(-s) \tag{26}$$

or, since $\Psi(s) = I$,

$$\Phi(s) = F^{-1}(s)\,G(s)\,G'(-s)\,F'^{-1}\,(-s). \tag{27}$$

If we define the polynomial $p(s)$ by

$$p(s) = |F(s)| = \prod_{r=1}^{m} (s-\lambda_r) \tag{28}$$

* In engineering literature the variables $\zeta(t)$ are called independent white noise sources. See Laning & Battin (1956), pp. 136–44 and Miller (1956), pp. 282–4.

and write $F^a(s)$ for the adjoint of $F(s)$, (27) becomes

$$\Phi(s) = \frac{F^a(s)\,G(s)\,G'(-s)\,F'^a(-s)}{p(s)\,p(-s)}. \tag{29}$$

For simplicity we shall consider in this paper only cases in which the roots $\lambda_r, r = 1, 2, \ldots, m$, of the equation $p(s) = 0$ are all distinct, and shall also assume that $p(s)$ does not contain any factor which is common to all the elements of $F^a(s)$. Expanding the right-hand side of (29) in partial fractions we then obtain

$$\Phi(s) = \sum_{r=1}^{m} \frac{K_r}{s-\lambda_r} + \sum_{r=1}^{m} \frac{K'_r}{-s-\lambda_r}, \tag{30}$$

where

$$K_r = \left.\frac{(s-\lambda_r)\,F^a(s)\,G(s)\,G'(-s)\,F'^a(-s)}{p(s)\,p(-s)}\right|_{s=\lambda_r}. \tag{31}$$

For a simple root λ_r the matrix $F(\lambda_r)$ is necessarily simply degenerate. The adjoint matrix $F^a(\lambda_r)$ is of unit rank and can be written

$$F^a(\lambda_r) = k_r \kappa_r, \tag{32}$$

where k_r is a column vector which satisfies the equation

$$F(\lambda_r)\,k_r = 0 \tag{33}$$

and κ_r is a row vector which satisfies

$$\kappa_r F(\lambda_r) = 0. \tag{34}$$

(See Frazer, Duncan & Collar (1950) 5·6.) Let l_r be a row vector defined by

$$l_r = \left.\frac{(s-\lambda_r)\,\kappa_r G(s)\,G'(-s)\,F'^a(-s)}{p(s)\,p(-s)}\right|_{s=\lambda_r}. \tag{35}$$

Then substituting (32) in (31) and using (35) we have

$$K_r = k_r l_r. \tag{36}$$

Thus the matrix K_r is of unit rank, with columns proportional to the vector k_r. It follows from (36) and (33) that

$$F(\lambda_r)\,K_r = 0 \quad (r = 1, 2, \ldots, m), \tag{37}$$

a result which could also have been obtained by premultiplying (31) by $F(s)$ and putting $s = \lambda_r$.

The matrix $\phi(\tau)$ of the system covariance functions is obtained by taking the inverse transform of (30). Formally, (30) does not have a unique inverse. However, for a stationary system the roots λ_r are negative or have negative real parts and $\phi(\tau) \to 0$ as $\tau \to \pm\infty$. It can readily be verified that the only inverse which is consistent with these conditions is that obtained by taking the first sum in (30) to be the one-sided transform of $\phi(\tau)$ for $\tau \geqslant 0$ and the second sum to be the one-sided transform for $\tau < 0$. The inverse of the first sum in (30) then gives

$$\phi(\tau) = \sum_{r=1}^{m} K_r e^{\lambda_r \tau} \quad (\tau \geqslant 0) \tag{38}$$

while from the inverse of the second sum we see that $\phi(-\tau) = \phi'(\tau)$, a fact which is also obvious from (25).

5. Calculation of $\Phi(s)$ from given points on $\phi(\tau)$

Let ϕ_τ be the matrix of discrete functions consisting of points at unit intervals of τ on the continuous functions $\phi(\tau)$, i.e. let

$$\phi_\tau = \phi(\tau) \quad \text{for} \quad (\tau = \ldots, -2, -1, 0, 1, 2, \ldots), \tag{39}$$

and let $M(z)$ be the one-sided generating function defined by

$$M(z) = \sum_{\tau=0}^{\infty} \phi_\tau z^\tau. \tag{40}$$

Then from (38), (39) and (40) we find that

$$M(z) = \sum_{r=1}^{m} \frac{K_r}{1-\mu_r z}, \tag{41}$$

where

$$\mu_r = e^{\lambda_r}. \tag{42}$$

We define the polynomial

$$\Delta(z) = \prod_{r=1}^{m} (1-\mu_r z), \tag{43}$$

and seek matrices $U(z)$ and $V(z)$, whose elements are polynomials in z, such that

$$M(z) = U^{-1}(z)\, V(z). \tag{44}$$

If (44) is to hold we require that

$$|U(z)| = \Delta(z) \tag{45}$$

and that

$$U(z) \sum_{r=1}^{m} \frac{K_r}{1-\mu_r z} = V(z). \tag{46}$$

Multiplying (46) by $1-\mu_q z$ we obtain

$$U(z)\left\{K_q + (1-\mu_q z) \sum_{r \neq q} \frac{K_r}{1-\mu_r z}\right\} = (1-\mu_q z)\, V(z) \tag{47}$$

and on putting $z = 1/\mu_q$, $q = 1, 2, \ldots, m$, this gives

$$U\left(\frac{1}{\mu_q}\right) K_q = 0 \quad (q = 1, 2, \ldots, m). \tag{48}$$

Since K_q is of unit rank, (48) provides m scalar equations in the coefficients of the polynomials occurring in each row of $U(z)$. If $U(z)$ is given the form

$$U(z) = I + U_1 z + U_2 z^2 + \ldots + U_c z^c, \tag{49}$$

where $U_1, U_2, \ldots, U_c$ are $n \times n$ matrices whose elements are constants, each row of $U(z)$ will contain cn coefficients of powers of z (excluding the constant terms, which are unity or zero). It can be shown that if c is given a suitable value and $cn-m$ coefficients in each row of $U(z)$ are put equal to zero, equations (48) suffice to determine the remaining m coefficients in each row of $U(z)$, provided the positions of the zeros are chosen in such a way that $|U(z)|$ is of degree m. It will always be possible to find positions for the zeros such that this condition is satisfied, and equation (45) will then also be satisfied. If m is an integral multiple of n it will usually, but not always, be possible to take c equal to m/n, in which case $cn-m = 0$ and none of the coefficients is put equal to zero.

When $U(z)$ has been determined $V(z)$ can be found from (46). It can be seen from (48) that $1-\mu_r z$ is a factor of $U(z)\,K_r$ and it follows from this and (46) that the degree of $V(z)$ is at least one less than that of $U(z)$. Thus $V(z)$ is of the form

$$V(z) = V_0 + V_1 z + \ldots + V_{c-1} z^{c-1}, \tag{50}$$

where $V_0, V_1, \ldots, V_{-1}$ are matrices whose elements are constants.

Premultiplying (44) by $U(z)$ and writing the matrices in expanded form, we have

$$(I + U_1 z + \ldots + U_c z^c)(\phi_0 + \phi_1 z + \ldots) = V_0 + V_1 z + \ldots + V_{c-1} z^{c-1}. \tag{51}$$

Equating the coefficients of $z^c, z^{c+1}, \ldots$, in (51) we obtain the difference equation

$$\phi_\tau + U_1 \phi_{\tau-1} + \ldots + U_c \phi_{\tau-c} = 0 \quad (\tau \geqslant c), \tag{52}$$

from which the matrices $U_1, U_2, \ldots, U_c$ can be calculated if ϕ_τ is given for a sufficient number of values of τ. Equating the constant terms in (51) we have

$$V_0 = \phi_0 \tag{53}$$

and equating the coefficients of $z, z^2, \ldots, z^{c-1}$ we find

$$V_l = \phi_l + \sum_{k=1}^{l} U_k \phi_{l-k} \quad (l = 1, 2, \ldots, c-1), \tag{54}$$

from which the matrices $V_0, V_1, \ldots, V_{c-1}$ can be calculated.

Thus given a sufficient number of points at equal intervals on the system covariance functions $\phi(\tau)$ it is possible to calculate $U(z)$ and $V(z)$. $M(z)$ is then given by (44), and can be expanded in partial fractions to give (41), from which the matrices K_r and the scalars μ_r are obtained immediately. The roots λ_r of the equation $p(s) = 0$ can then be calculated from

$$\lambda_r = \log \mu_r, \tag{55}$$

which follows from (42), and the matrix $\Phi(s)$ can then be found from the K_r and λ_r using (30).

6. Calculation of $Q(s)$ and $\theta(s)$ from $\Phi(s)$ when all equations are of the same order

Consider an n variable system defined by (20) or (23), in which all the scalar equations of $F(D)\,x(t) = G(D)\,\zeta(t)$ are of the same order u, and which cannot be reduced to an equivalent system in which any equation is of lower order. The elements on the leading diagonal of $F(s)$ are polynomials of degree u; but as a result of the restriction on the degree of the numerator terms of the transfer functions in the matrix $Q(s)$, discussed in § 2 above, the non-diagonal elements are of degree $u-1$ at most. The coefficient of s^u in each of the diagonal elements of $F(s)$ can be taken as unity. There are u other coefficients of powers of s (including the constant term) in each of the diagonal elements and u coefficients in each of the non-diagonal elements. In calculating $F(s)$ from $\Phi(s)$, or equivalently from the K_r and λ_r, $r = 1, 2, \ldots, m$, there are therefore nu coefficients to be determined in each row of $F(s)$. Since $|F(s)|$ is of degree nu we have $m = nu$. We obtain from (37) a set of nu non-homogeneous equations in the coefficients of each row of $F(s)$. It can be shown that the equations in each set are linearly independent. Thus all the coefficients in $F(s)$ are uniquely determined.

Given $F(s)$ we can find $Q(s)$ from

$$Q(s) = P^{-1}(s)\,F(s), \tag{56}$$

which follows from (21). It will be noticed that the elements of the diagonal matrix $P(s)$ are identical with the elements on the leading diagonal of $F(s)$. The numerator and denominator polynomials of the elements of $Q(s)$ as calculated from (56) may have common factors which can be cancelled.

Given $F(s)$ and the K_r and λ_r, $G(s)\,G'(-s)$ is readily obtained from the equation

$$G(s)\,G'(-s) = \sum_{r=1}^{m} \frac{F(s)\,K_r F'(-s)}{s-\lambda_r} + \sum_{r=1}^{m} \frac{F(s)\,K_r' F'(-s)}{-s-\lambda_r} \tag{57}$$

which is derived from (27) and (30). Using (19) and (22) we then obtain

$$\theta(s) = P^{-1}(s)\,G(s)\,G'(-s)\,P'^{-1}(-s). \tag{58}$$

The numerator and denominator polynomials of the elements of $\theta(s)$ as calculated from (58) may also have common factors which can be cancelled.

7. Calculation of $Q(s)$ and $\theta(s)$ from $\Phi(s)$ when all the equations are not of the same order, but the disturbances are not cross-correlated

If all the equations in a system are not of the same order they cannot all be identified unless further restrictions are placed on the system. Consider a system $F(D)\,x(t) = G(D)\,\zeta(t)$ in which the equations are not all of the same order, and in which the arrangement of the equations is such that the first equation is of the lowest order and the order of each successive equation is at least as high as that of the one preceding it. If (27) is satisfied by the matrices $F(s)$ and $G(s)$ of this system it is also satisfied by $\mu F(s)$ and $\mu G(s)$ where μ is an arbitrary matrix of constants. We can add to any row of $F(s)$ a multiple of an earlier row corresponding to an equation of lower order without violating the condition that the non-diagonal elements in each row be of lower degree than the diagonal elements; but we cannot add a multiple of any succeeding row without violating this condition. As a result of this condition, therefore, μ is restricted to a triangular matrix with zeros above the leading diagonal. Given $\Phi(s)$ for this system, the first equation is therefore determinate (and so are any other equations of the same order as the first one) but the remaining equations of higher order than the first one are not identified without further restrictions on the system.

If it is known that the disturbances to the system are not cross-correlated, $G(s)\,G'(-s)$ must be diagonal. If $\mu F(s)$ and $\mu G(s)$ are to be admissible solutions, therefore, μ must satisfy the further restriction that the non-diagonal elements of $\mu G(s)\,G'(-s)\,\mu'$ be identically zero. Writing g_{ii} for the diagonal elements of $G(s)\,G'(-s)$ the first part of the array formed by the elements lying above the leading diagonal of the matrix product $\mu G(s)\,G'(-s)\,\mu'$ is found to be

$$\begin{array}{lll} \mu_{11}g_{11}\mu_{21} & \mu_{11}g_{11}\mu_{31} & \mu_{11}g_{11}\mu_{41} \\ & \mu_{21}g_{11}\mu_{31}+\mu_{22}g_{22}\mu_{32} & \mu_{21}g_{11}\mu_{41}+\mu_{22}g_{22}\mu_{42} \\ & & \mu_{31}g_{11}\mu_{41}+\mu_{32}g_{22}\mu_{42}+\mu_{33}g_{33}\mu_{43}. \end{array}$$

Since u_{ii} cannot be made zero, we see from the first element in the array that if $g_{11} \neq 0$, μ_{21} must vanish. From the second row of the array we see that if $g_{11} \neq 0$ and $g_{22} \neq 0$, μ_{31} and μ_{32} must vanish. Similarly, from the third row, μ_{41}, μ_{42} and μ_{43} must vanish if g_{11}, g_{22} and g_{33}

are non-zero. Continuing in this way it can be shown that all the elements μ_{ij}, $(i > j)$ must vanish if all the elements g_{ii}, $(i = 1, 2, \ldots, n-1)$ are non-zero. The matrix μ is then diagonal, and all the equations in the system are identified.

To calculate $F(s)$ for the system we have just been considering, given $\Phi(s)$ we first note that since the equations of lowest order are identified without restriction of the disturbance covariance functions, it will be possible to find coefficients in at least one of the rows of $F(s)$ which satisfy the appropriate set of scalar equations derived from (37) for all the value of r. The number of equations will be greater than the number of non-zero coefficients in this row, but they will form a consistent set with a sufficient number of independent equations to determine the non-zero coefficients. In the present case the first equation in the system is of lowest order, so the coefficients in the first row of $F(s)$ can be found in this way.

From (27) we have

$$F(s)\,\Phi(s)\,F'(-s) = G(s)\,G'(-s). \tag{59}$$

On the assumption that $G(s)\,G'(-s)$ is diagonal the first column of the matrix product in (59) yields the equations

$$\sum_{j=1}^{n}\sum_{q=1}^{n} F_{ij}(s)\,\Phi_{jq}(s)F_{1q}(-s) = 0 \quad (i = 2, 3, \ldots, n). \tag{60}$$

If we give $\Phi_{jq}(s)$ and $F_{1q}(-s)$ in (60) their known numerical values, multiply through by $p(s)\,p(-s)$ and equate the coefficients of each power of s we obtain for each value of i a consistent set of linear equations in the coefficients of the polynomials in the ith row of $F(s)$, which can be solved for these coefficients. $Q(s)$ can then be calculated from $F(s)$, and $\theta(s)$ is given by

$$\theta(s) = Q(s)\,\Phi(s)\,Q'(-s). \tag{61}$$

8. Stochastic difference equation for discrete points on $x(t)$

Let $\Gamma(z)$ be the two-sided generating function of the discrete functions ϕ_τ defined in (39), so that

$$\Gamma(z) = \sum_{\tau=-\infty}^{\infty} \phi_\tau z^\tau. \tag{62}$$

Since $\phi_{-\tau} = \phi'_\tau$ we have from (40)

$$\Gamma(z) = M(z) + M'(z^{-1}) - \phi_0. \tag{63}$$

Substituting (44) in (63) we obtain

$$\Gamma(z) = U^{-1}(z)\,V(z) + V'(z^{-1})\,U'^{-1}(z^{-1}) - \phi_0 \tag{64}$$

or

$$U(z)\,\Gamma(z)\,U'(z^{-1}) = W(z)\,W'(z^{-1}), \tag{65}$$

where $W(z)$ is any matrix which satisfies the equation

$$W(z)\,W'(z^{-1}) = V(z)\,U'(z^{-1}) + U(z)\,V'(z^{-1}) - U(z)\,\phi_0\,U'(z^{-1}). \tag{66}$$

It follows from (65) that $\Gamma(z)$ is also the generating function of the covariance functions of discrete variables x_t which are generated by the system of difference equations

$$U(E^{-1})\,x_t = W(E^{-1})\,\epsilon_t \tag{67}$$

from stationary stochastic variables ϵ_t with covariance functions

$$E\{\epsilon_t \epsilon'_{t-\tau}\} = \begin{cases} I & \text{if} \quad \tau = 0, \\ 0 & \text{if} \quad \tau \neq 0. \end{cases} \tag{68}$$

(The symbol E^{-1} in (67) is the shift operator, defined by $E^{-h}y_t = y_{t-h}$.)

The discrete variables x_t in (67) may be identified with points at unit time intervals on the continuous variables $x(t)$ in (20), and it follows that the sampling properties of estimates of the continuous system (20) may be studied by considering the sampling properties of estimates of the discrete system (67). The work done by Bartlett (1946, 1955), M. G. Kendall (1945, 1946, 1949), Quenouille (1947, 1957), Whittle (1952, 1953), Wold (1949) and others on sampling properties and goodness-of-fit tests for autoregressive, moving average and mixed processes is therefore relevant to the estimation of parameters in systems of stochastic differential equations. Most of the theoretical results which have been obtained, however, are valid only for large samples.

Some sampling experiments are being carried out on systems of the type considered in this paper. The results will be published later.

REFERENCES

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time series. *J. Roy. Statist. Soc.* B, **8**, 27.

Bartlett, M. S. (1955). *Stochastic Processes*. Cambridge University Press.

Frazer, R. A., Duncan, W. J. & Collar, A. R. (1950). *Elementary Matrices*. Cambridge University Press.

James, H. M., Nichols, N. B. & Phillips, R. S. (1947). *Theory of Servomechanisms*. New York: McGraw Hill.

Kendall, M. G. (1945). On the analysis of oscillatory time series. *J. Roy. Statist. Soc.* **108**, 93.

Kendall, M. G. (1946). *Contributions to the Study of Oscillatory Time Series*. Cambridge University Press.

Kendall, M. G. (1949). Tables of autoregressive series. *Biometrika*, **36**, 267.

Laning, J. H. & Battin, R. H. (1956). *Random Processes in Automatic Control*. New York: McGraw Hill.

Miller, K. S. (1956). *Engineering Mathematics*. London: Constable and Co. Ltd.

Quenouille, M. H. (1947). A large-sample test for the goodness-of-fit of autoregressive schemes. *J. Roy. Statist. Soc.* A, **110**, 123.

Quenouille, M. H. (1957). *The Analysis of Multiple Time Series*. London: Charles Griffin and Co. Ltd.

Whittle, P. (1952). Tests of fit in time series. *Biometrika*, **39**, 309.

Whittle, P. (1953). The analysis of multiple stationary time-series. *J. Roy. Statist. Soc.* B, **15**, 125.

Widder, D. V. (1946). *The Laplace Transform*. Princeton University Press.

Wold, H. (1949). A large-sample test for moving averages. *J. Roy. Statist. Soc.* **11**, 297.

DESIGN OF EXPERIMENTS IN NON-LINEAR SITUATIONS

By G. E. P. BOX and H. L. LUCAS

Statistical Techniques Research Group, Princeton University and Institute of Statistics, North Carolina State College

1. Outline of the problem

We suppose that some response η is a known function

$$\eta = f(\xi_1, \xi_2, \ldots, \xi_k; \theta_1, \theta_2, \ldots, \theta_p) = f(\boldsymbol{\xi}; \boldsymbol{\theta}) \tag{1}$$

of k variables whose levels are denoted by the elements $\xi_1, \ldots, \xi_i, \ldots, \xi_k$ of the vector $\boldsymbol{\xi}$ and of p parameters $\theta_1, \ldots, \theta_r, \ldots, \theta_p$ elements of the vector $\boldsymbol{\theta}$, and that this function is not necessarily linear in either the variables or the parameters. We further suppose that it is possible to perform a series of experiments or trials in each one of which the response is observed at a particular chosen set of levels of the variables decided in advance.

The problem here considered is that of selecting a programme of trials such that they may be expected to provide results from which the p parameters can be estimated with high accuracy. In general the experimental programme may be defined by an $N \times k$ matrix $\mathbf{D} = \{\xi_{iu}\}$ called the design matrix. The uth row $\boldsymbol{\xi}'_u$ of this matrix with elements $\xi_{1u}, \ldots, \xi_{iu}, \ldots, \xi_{ku}$ provides the levels of the k variables at which the response is to be observed in the uth trial.

1·1. *Example*

The following example illustrates the situation we have in mind. Suppose a consecutive chemical reaction is under study in which a substance A decomposes to form substance B which then in turn decomposes to form substance C. With the assumptions that the reactions are first order and irreversible it may be shown that, after time ξ_1 has elapsed, the yield η of intermediate product B is given by

$$\eta = \frac{\theta_1}{\theta_1 - \theta_2}\{\exp(-\theta_2 \xi_1) - \exp(-\theta_1 \xi_1)\}, \tag{2}$$

where θ_1 and θ_2 are constants measuring the specific rates of the first and second decompositions, respectively.

The problem is to choose a set of times $\xi_{1u}, u = 1, 2, \ldots, N$, at which to observe the yield so that from these observations θ_1 and θ_2 can be estimated as accurately as possible, allowing for experimental error in the observations. In this example the design matrix $\mathbf{D}$ would simply consist of a single column whose N elements were the times $\xi_{11}, \ldots, \xi_{1u}, \ldots, \xi_{1N}$ at which the yield was to be observed and $\boldsymbol{\theta}$ would contain the two elements θ_1 and θ_2. More generally there will be k variables (for example, ξ_1 = time, ξ_2 = temperature, ξ_3 = concentration, etc.) and the design matrix will contain k columns.

1·2. *Practical limits to levels of variables*

In practice, levels at which variables such as time, temperature, concentration and dosage can be set are usually restricted. For example, in the study of the chemical reaction above, the feasible range for the variable time (ξ_1) will be from zero to some value ξ_1(max) which will be the longest time which could be contemplated in an actual experiment. In general, there

will exist some experimental region R in the $\boldsymbol{\xi}$ space which delimits the area within which trials can actually be conducted. In a chemical reaction, for example, such limitation may occur from explosive hazards associated with certain combinations of temperature and pressure. In other cases it may be simply that it is experimentally inconvenient to conduct trials at certain combinations of conditions. In some cases the experimental region can be defined by a series of inequalities in the ξ's of the type

$$\xi_i(\min) \leqslant \xi_i \leqslant \xi_i(\max) \quad (i = 1, 2, \ldots, k).$$

In general, however, the definition of R may be more complicated.

1·3. *Criterion for selection of design*

We denote the response observed at the uth set of experimental conditions by y_u and we suppose that

$$E(y_u) = \eta_u = f(\boldsymbol{\xi}_u, \boldsymbol{\theta}), \tag{3}$$

and

$$E(y_u - \eta_u)(y_v - \eta_v) = \begin{cases} \sigma^2 & u = v \\ 0 & u \neq v \end{cases} \quad (u, v = 1, 2, \ldots, N), \tag{4}$$

where, in general, σ^2 is unknown. The true values of the parameters are denoted by $\theta_{10}, \theta_{20}, \ldots, \theta_{p0}$, the elements of the vector $\boldsymbol{\theta}_0$. The partial derivatives of the response function with respect to the rth parameter θ_r for the uth set of experimental conditions, taken at the point $\boldsymbol{\theta}_0$, is denoted by

$$f_{ru} = \left[\frac{\partial f(\boldsymbol{\xi}, \boldsymbol{\theta})}{\partial \theta_r}\right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}. \tag{5}$$

Finally, the $N \times p$ matrix of these derivatives is

$$\mathbf{F} = \{f_{ru}\}. \tag{6}$$

It is well known that the least squares estimates $\hat{\boldsymbol{\theta}}$ obtained by minimizing the sum of squares

$$\sum_{u=1}^{N} \{y_u - f(\boldsymbol{\xi}_u, \boldsymbol{\theta})\}^2 \tag{7}$$

and given by solving the normal equations

$$\sum_{u=1}^{N} \{y_u - f(\boldsymbol{\xi}_u, \hat{\boldsymbol{\theta}})\} \{\partial f(\boldsymbol{\xi}_u, \boldsymbol{\theta})/\partial \theta_r\}_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0 \quad (r = 1, 2, \ldots, p) \tag{8}$$

have a variance covariance matrix which is approximated by $(\mathbf{F}'\,\mathbf{F})^{-1}\sigma^2$.

We shall proceed by attempting to choose $\mathbf{D}$ so that the determinant $|(\mathbf{F}'\mathbf{F})^{-1}|$ is made as small as possible. If we assume that each $f(\boldsymbol{\xi}_u, \boldsymbol{\theta})$ is approximately linear in the neighbourhood of $\boldsymbol{\theta}_0$ we are thus minimizing Wilk's generalized variance for the estimates. The logic of this choice, also made by Wald (1943), can be seen by noting that if the estimates are approximately normally distributed, the determinant $|(\mathbf{F}'\,\mathbf{F})^{-1}|$ is proportional to the volume contained within any specific ellipsoidal probability contour for $\boldsymbol{\theta}$ about $\boldsymbol{\theta}_0$ in the space of the parameters. The criterion chosen ensures that any such probability contour includes the smallest volume.

To understand the significance of this criterion in general non-linear situations it seems desirable to consider the relationship between sample space, the 'solution locus' within sample space, and parameter space. The point $\boldsymbol{\theta}$ in parameter space can be mapped on to the point $P(\boldsymbol{\theta})$ in sample space with co-ordinates $\eta_u = f(\boldsymbol{\xi}_u, \boldsymbol{\theta})$ $(u = 1, \ldots, N)$. We define the part of sample space covered by this mapping as the 'solution locus'.

Now, given any point $P(\boldsymbol{\theta})$ on the solution locus, we can make an orthogonal transformation $\boldsymbol{\eta} = \mathbf{H}\boldsymbol{\zeta}$ of co-ordinates in sample space such that the solution locus in the immediate neighbourhood of $P(\boldsymbol{\theta})$ coincides with part of the p-dimensional hyperplane $\zeta_i = \text{constant}$ $(i = p+1, \ldots, N)$. We can then interpret $|(\mathbf{F}'\mathbf{F})^{-1}|$ as the square of the Jacobian of the transformation from co-ordinates $(\zeta_1, \ldots, \zeta_p)$ in sample space to $(\theta_1, \ldots, \theta_p)$ in parameter space, when $P(\boldsymbol{\theta}) = P(\boldsymbol{\theta}_0)$. So the criterion that $|(\mathbf{F}'\mathbf{F})^{-1}|$ is to be made as small as possible means geometrically that the p-dimensional volume of that part of parameter space associated with a region of prescribed small volume round $P(\boldsymbol{\theta}_0)$ on the solution locus shall be as small as possible. Since we do not know $\boldsymbol{\theta}_0$ in advance, we must assume that we can find a design that makes this Jacobian about as small as possible for all points $P(\boldsymbol{\theta})$ near $P(\boldsymbol{\theta}_0)$.

Let us now assume that the part of the solution locus near $P(\boldsymbol{\theta}_0)$ can be approximated by a p-dimensional linear subspace of sample space. If we also assume that the experimental error distributions are normally distributed as well as satisfying (3) and (4), then the confidence region for $\boldsymbol{\theta}$ based on the likelihood ratio criterion consists of all points associated with points on the solution locus within a p-dimensional sphere centred on the least squares estimate $P(\hat{\boldsymbol{\theta}})$, and with radius independent of the experimental design. Any design that gives an adequate approximation to the smallest possible value of the Jacobian for all points round $P(\boldsymbol{\theta}_0)$ therefore approximately minimizes the p-dimensional volume of the confidence region in parameter space.

1·4. *Necessity for preliminary estimates*

We see that the efficiency of different possible designs depends upon the matrix $\mathbf{F}$ whose elements are the values of the derivatives of the response function with respect to the θ's at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Now the derivatives $\partial f/\partial\theta_r$ can only be independent of the values actually taken by the parameter θ_r if the response function f is linear in θ_r. For non-linear response functions the values of the derivatives, and hence the efficiency of any particular design, depend upon the actual values of the p-parameters. If we are to suppose that effective design is possible at all, we must also assume therefore that *something* is known about the values of the parameters in advance.

In practical problems it will almost invariably be the case that some such information is available, and this will then provide the basis of a first design. As is inevitably the case here and elsewhere, the real effectiveness of the design will depend upon the reliability of the information on which it is based. In this paper, therefore, we assume that preliminary values $\boldsymbol{\theta}^*$ of the parameters are available and will proceed so that the design would be optimal if these quantities were the true values. Some further discussion of the general situation for which this theory is appropriate is given by Box (1957).

1·5. *Possible extensions of the problem*

It is clear that in some examples normality of distribution and constancy of variance would not be possible at all points on the response function if only because the variation might be limited either upwards or downwards. Two obvious extensions, therefore, of this work would be to generalize to a maximum likelihood solution in which any distribution is considered or to consider weighted least squares. Valuable work in this direction has already been carried out by Elfving (1952) and Chernoff (1953). Again, in some cases we might wish to make a transformation of the model before the present theory was applied.

Apart from questions of appropriate distribution assumptions it should be possible to extend this work in at least two other directions. In the first place in different problems the reliability of the preliminary values $\boldsymbol{\theta}^*$ will be different, and we might take account of this by using some rough estimate of the prior probability distribution. Again, some observations might be cheaper than others. One could then attempt to find the design giving the most information for a given cost. Finally, it would be desirable in many circumstances to use a sequential procedure in which estimates from a preliminary experiment would be used to plan a second experiment, the results of which would in turn be used to plan a third experiment, and so on. These possibilities will not be further discussed here but we hope it may be possible to pursue them in later work.

2. Solution

Suppose that we are given the functional form

$$\eta_u = f(\boldsymbol{\xi}_u, \boldsymbol{\theta}) \tag{9}$$

and a set of preliminary values $\boldsymbol{\theta}^*$. Then we shall choose the design $\mathbf{D}$ such that $|\mathbf{F}^{*\prime}\mathbf{F}^*|$ is a maximum within the experimental region R in the $\boldsymbol{\xi}$-space, where

$$\mathbf{F}^* = \{f^*_{ru}\} \tag{10}$$

and

$$f^*_{ru} = \left[\frac{\partial f(\boldsymbol{\xi}_u, \boldsymbol{\theta})}{\partial \theta_r}\right]_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}. \tag{11}$$

To limit to manageable proportions the possibilities we need to consider, we will in this paper assume that the number of trials (that is the number of combinations of levels of the variables) is equal to the number p of parameters to be estimated. Each of the p trials could of course be replicated to attain any desired degree of precision, but it should be noted that an r-fold replication of a design that minimizes the generalized variance using p observations, does not necessarily minimize the generalized variance for all possible experiments involving rp observations. Experiments in which the number of parameters estimated is equal to the number of trials are not perhaps intuitively appealing. This is (i) because one is usually not completely assured that the model is correct; so that one ordinarily has in fact a twofold objective: to check the model and to estimate the parameters in it if it proves to be representationally adequate, and (ii) because designs in which there are more than the minimum number of points may provide greater protection against gross errors in guessing $\boldsymbol{\theta}^*$.

In order for a design to be efficient in 'checking' the model, we must be specific about what sorts of departure from the model we suspect. In so far as this can be done, we should be able to write an 'alternative' wider model which makes allowance for these possible departures *in terms of extra parameters to be estimated*. In this way, within the framework discussed we can achieve both objectives mentioned above. As a simple illustration consider the chemical reaction discussed earlier. Here A was decomposed to form B which in turn decomposed to C. In a particular problem the second decomposition might be unlikely, in which case the primary model in mind for the production of B would be

$$\eta = 1 - e^{-\theta_1 \xi}. \tag{12}$$

However, in the light of possible decomposition of B to C it would be appropriate to plan the experiment so as to be effective in estimating both θ_1 and θ_2 in the model of equation (2).

When nothing is known about the type of departures which may occur we are confronted once more with the fact already mentioned that the effectiveness of a design must depend on

the reliability of information on which it is based. If this information is non-existent then we have no basis on which to plan. In practice, however, the situation may arise where some information is available which is difficult to express in terms of extra parameters in the model, but is manifested in feelings about the way the response surface may be distorted. In this case further points might be added to the design in an intuitive way, so as to tie down doubtful areas of the surface, while at the same time bearing in mind the locations already determined which ensure maximum reliability for the principal parameters when the model is adequate.

When gross errors may occur in guessing $\boldsymbol{\theta}^*$ the present theory can be used to explore the changes that would occur in the position of the optimal p design points with changes in $\boldsymbol{\theta}^*$ over the region of uncertainty and the losses in efficiency which occurred with given errors in $\boldsymbol{\theta}^*$. If more than p-points were available these might then be chosen so as to minimize the loss of efficiency as far as possible.

2·2. *A geometrical formulation*

For that class of designs for which the number of trials is equal to the number of parameters, $\mathbf{F}^*$ is a square $(p \times p)$ matrix. Consequently,

$$|\mathbf{F}^{*\prime}\mathbf{F}^*| = |\mathbf{F}^*|^2 \tag{13}$$

and we should choose the design $\mathbf{D}$ so that the modulus of $|\mathbf{F}^*|$ is a maximum in the experimental region R.

In understanding the problem to be solved and in showing the properties of solutions in simple cases, the following geometrical formulation has been found of help. Consider a p-dimensional space in which the co-ordinates are the derivatives

$$\left(\frac{\partial f(\boldsymbol{\xi}, \boldsymbol{\theta})}{\partial \theta_r}\right)_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} = \phi_r(\boldsymbol{\xi}) \quad (r = 1, 2, \ldots, p), \tag{14}$$

where, it will be noted, $\phi_r(\boldsymbol{\xi})$ is a function of $\boldsymbol{\xi}$ only. The surface in the space of the derivatives generated by all permissible values of $\boldsymbol{\xi}$ we call the 'design locus'. Thus with any particular choice of p 'design points', $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_u, \ldots, \boldsymbol{\xi}_p$, in the ξ-space for a particular design matrix $\mathbf{D}$, there will be corresponding points on the design locus in the space of the derivatives $\phi_r(\boldsymbol{\xi})$.

Now the volume of a simplex in p-space formed by p-points

$$(z_{11}, z_{12}, \ldots, z_{1p}); \ldots; (z_{s1}, z_{s2}, \ldots, z_{sp}); \ldots; (z_{p1}, z_{p2}, \ldots, z_{pp})$$

and the origin $(0, 0, \ldots, 0)$ is proportional to the determinant of the $p \times p$ matrix $\{z_{sr}\}$. Consequently, we must find p design points $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_u, \ldots, \boldsymbol{\xi}_p$ for which the corresponding points in the space of the derivatives, with co-ordinates defined by equation (14), form with the origin the simplex of greatest volume.

2·3. *Example*

As a simple illustration, consider the example given in § 1·1 of a consecutive first-order chemical reaction. Suppose that we are given preliminary guesses θ_1^*, θ_2^* for the values of the parameters. We wish to choose two values ξ_{11}, ξ_{12} of time so that the best estimates will be available for θ_1, θ_2. We have

$$\eta = f(\xi_1; \theta_1, \theta_2) = \frac{\theta_1}{\theta_1 - \theta_2}\{\exp(-\theta_2 \xi_1) - \exp(-\theta_1 \xi_1)\}. \tag{15}$$

The region R is defined by $0 \leqslant \xi_{11} \leqslant \xi_{12} \leqslant \xi_1(\max)$. (16)

but for the moment no upper limit $\xi_1(\max)$ will be imposed. We have to choose ξ_{11}, ξ_{12} so as to maximize the modulus of the determinant

$$\Delta = |\mathbf{F}^*| = \begin{vmatrix} \phi_1(\xi_{11}) & \phi_2(\xi_{11}) \\ \phi_1(\xi_{12}) & \phi_2(\xi_{12}) \end{vmatrix}. \tag{17}$$

If the preliminary guesses are

$$\theta_1^* = 0{\cdot}7, \quad \theta_2^* = 0{\cdot}2, \tag{18}$$

then
$$\phi_1(\xi_1) = (0{\cdot}8 + 1{\cdot}4\xi_1)\exp(-0{\cdot}7\xi_1) - 0{\cdot}8\exp(-0{\cdot}2\xi_1), \tag{19}$$

$$\phi_2(\xi_1) = (2{\cdot}8 - 1{\cdot}4\xi_1)\exp(-0{\cdot}2\xi_1) - 2{\cdot}8\exp(-0{\cdot}7\xi_1). \tag{20}$$

We can of course substitute equations (19) and (20) in (17) and calculate the values ξ_{11}, ξ_{12} which maximize the modulus of Δ. It is instructive, however, to study this solution in some detail using the geometrical construction. The dotted line in Fig. 1 is a graph of the function

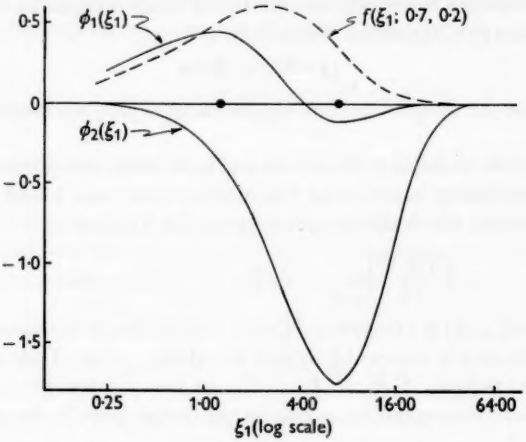


Fig. 1. The function $f(\xi_1; 0{\cdot}7, 0{\cdot}2)$ and its derivatives. The optimal solution for the design is indicated by the heavy dots.

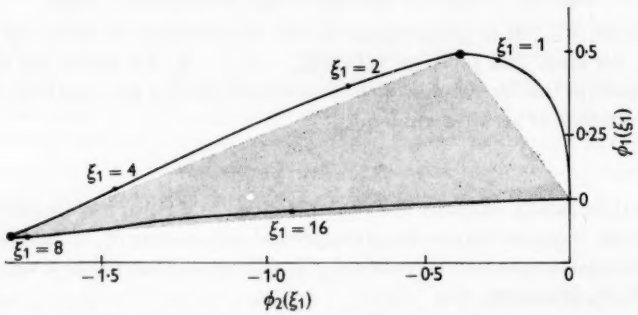


Fig. 2. The design locus for the function $f(\xi_1; 0{\cdot}7, 0{\cdot}2)$ with the optimal solution indicated by heavy dots.

$f(\xi_1; 0{\cdot}7, 0{\cdot}2)$. On the same diagram are shown the values of the derivatives $\phi_1(\xi_1)$, $\phi_2(\xi_1)$. Fig. 2 shows $\phi_1(\xi_1)$ plotted against $\phi_2(\xi_1)$ for values of ξ_1 covering the range of interest. This is the curve we call the design locus.

In terms of Fig. 2 the problem is to find those two points ξ_{11}, ξ_{12} on this locus, which together with the origin form the vertices of the triangle of greatest area. The exact solution obtained by numerical methods described later is $\xi_{11} = 1{\cdot}23$, $\xi_{12} = 6{\cdot}86$, corresponding to the triangle denoted by the shaded area in Fig. 2. This solution is indicated in Fig. 1 by the points at $\xi_{11} = 1{\cdot}23$ and $\xi_{12} = 6{\cdot}86$. From the graph of the function $f(\xi_1; 0{\cdot}7, 0{\cdot}2)$ it will be seen that for this particular example the values are chosen so as to lie on either side of the maximum at points of moderately high yield. The graphs of the derivatives illustrate how the criterion seeks points at which the derivatives are large in absolute value, but also, so far as possible, uncorrelated.

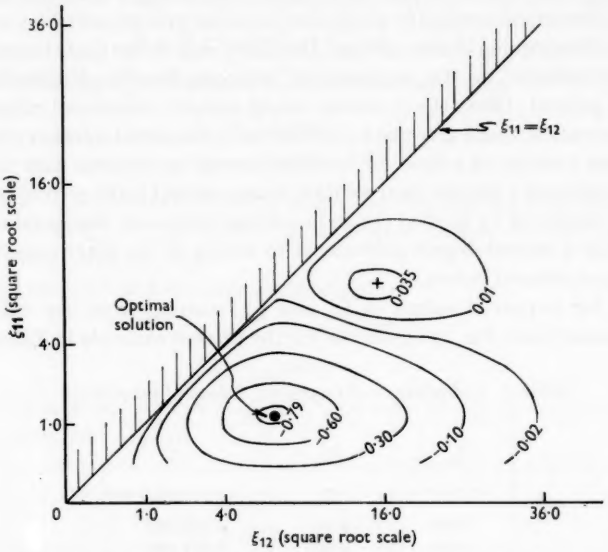


Fig. 3. Contours of Δ for the function $f(\xi_1; 0{\cdot}7, 0{\cdot}2)$ in the space of ξ_{11}, ξ_{12} for $0 \leqslant \xi_{11} \leqslant \xi_{12}$.

2·4. *Some problems in determining the solution*

This example may be used to illustrate some of the problems in determining the solution. Fig. 3 shows contours of Δ in the space ξ_{11}, ξ_{12} as defined by equation (17) within the region $0 \leqslant \xi_{11} \leqslant \xi_{12}$. It is seen that this surface contains *two* local extrema, both corresponding to local maxima in the *modulus* of the determinant. In terms of the actual value of the determinant, there is a minimum of $\Delta = -0{\cdot}810$ at the values $\xi_{11} = 1{\cdot}23$, $\xi_{12} = 6{\cdot}86$ and a maximum of $\Delta = 0{\cdot}039$ at about $\xi_{11} = 7{\cdot}5$, $\xi_{12} = 15{\cdot}0$. The latter value is proportional to the area of a small triangle that can be fitted into the concavity on the lower side of the design locus in Fig. 2. This solution is of no practical interest, since the absolute value of the determinant is only about one-twentieth of the value which corresponds to the other solution. However, it is of some importance to note that surfaces with multiple extrema can occur, and may complicate the problem of finding best solutions in more complex cases.

2·5. *Numerical solution*

Although in particular examples there is an analytic solution to the problem of maximizing the modulus of the determinant $|\mathbf{F}^*|$, often numerical methods must be employed. Such methods are attractive because they allow us to obtain solutions even when the function $f(\boldsymbol{\xi}, \boldsymbol{\theta})$ is not given explicitly. In chemical examples, for instance, a kinetic mechanism can often be postulated in terms of a series of differential equations which may have no explicit solution. The equations may be integrated numerically, however, on a computer and the quantities $\phi_r(\boldsymbol{\xi})$ for any trial value $\boldsymbol{\xi}$ obtained from difference formulae. Iterative methods may then be employed to find the maximum value of the modulus of the determinant.

In the particular example studied it is seen from Fig. 3 that in the neighbourhood of the optimum solution the surface is well conditioned, containing no oblique ridges, and that consequently a numerical method in which the variables were adjusted one at a time would be successful in determining the maximum. It is fairly easy to see that the satisfactory state of conditioning indicated by the 'symmetrical' maxima found in this problem cannot be expected to be general. Often the Δ surface would contain attenuated ridges which in the later stages of iteration would give rise to difficulties in numerical solution when the method of adjusting 'one variable at a time' or 'steepest ascent' procedures were used. A method which can be employed when the final solution is approached is the second-order procedure similar to that employed by Koshal (1933) in solving maximum likelihood equations. This consists of fitting a second-degree polynomial to values of the determinant calculated at suitably arranged series of points.

Calculations for improved values of ξ_{11} and ξ_{12}, starting from the values $\xi_{11} = 1{\cdot}10$, $\xi_{12} = 6{\cdot}50$ (obtained from Fig. 2) are shown for the present example in Table 1.

Table 1. *Calculation of improved values of maximum*

ξ_{11}	ξ_{12}	$-\Delta$
1·10	6·50	0·803 336
1·20	6·50	0·807 992
1·30	6·50	0·806 709
1·10	7·00	0·804 758
1·10	7·50	0·798 920
1·20	7·00	0·809 632

Writing $x_1 = 5(\xi_{11} - 1{\cdot}10)$, $x_2 = (\xi_{12} - 6{\cdot}50)$, the second-degree interpolation polynomial is

$$-\Delta = 0{\cdot}803336 + 0{\cdot}015251x_1 + 0{\cdot}010104x_2 - 0{\cdot}011878x_1^2 - 0{\cdot}014250x_2^2 + 0{\cdot}000872x_1x_2$$

which has a maximum at $\quad x_1 = 0{\cdot}66, \quad x_2 = 0{\cdot}37,$

that is $\quad \xi_{11} = 1{\cdot}23, \quad \xi_{12} = 6{\cdot}87.$

A second iteration provided values in substantial agreement with these, namely, $\xi_{11} = 1{\cdot}23$ and $\xi_{12} = 6{\cdot}86$. It is clear that because of difficulties of possible multiple maxima and the complexities of the Δ surface, considerable caution is necessary in applying numerical

methods to problems in which there are a large number of parameters. It is hoped to discuss such cases in a later paper. The examples which follow contain only a few parameters but nevertheless are thought to be of some inherent interest.

3. Examples of non-linear designs

3·1. *Simple decay and growth law*

Consider the model

$$\eta = \theta_1 \exp(\theta_2 \xi_1), \quad 0 \leqslant \xi_1(\min) \leqslant \xi_{11} < \xi_{12} \leqslant \xi_1(\max). \tag{21}$$

With ξ representing time and with $\theta_1 > 0$, $\theta_2 < 0$, this is the well-known first-order decay law. With $\theta_1 > 0, \theta_2 > 0$, the function can represent growth phenomena. In application both lower and upper limits $\xi_1(\min)$ and $\xi_1(\max)$ may be imposed corresponding to the earliest and latest points in time at which observations can be made. The value of the determinant is

$$\Delta = \theta_1 \exp\{\theta_2(\xi_{11}+\xi_{12})\}(\xi_{12}-\xi_{11}). \tag{22}$$

The modulus of which for $\theta_2^* < 0$ attains its maximum value when

$$\left.\begin{aligned} \xi_{11} &= \xi_1(\min), \\ \xi_{12} &= \xi_1(\min) - 1/\theta_2^*. \end{aligned}\right\} \tag{23}$$

Substituting these values into equation (21) we see that the best design will be that in which a proportion $e^{-1} = 36{\cdot}8\,\%$ of the amount of substance present at the first observation remains at the second observation.

For $\theta_2^* > 0$ the modulus of Δ is a maximum when

$$\left.\begin{aligned} \xi_{11} &= \xi_1(\max) - 1/\theta_2^*, \\ \xi_{12} &= \xi_1(\max). \end{aligned}\right\} \tag{24}$$

The first observation should thus be made at a time when $36{\cdot}8\,\%$ of the growth expected at the final observation has occurred.

If instead of the amount of substance decayed we consider the amount of product formed from a simple decay law, we have

$$\eta = \theta_1\{1-\exp(\theta_2 \xi_1)\}$$

for

$$0 \leqslant \xi_1(\min) \leqslant \xi_{11} < \xi_{12} \leqslant \xi_1(\max). \tag{25}$$

The value of the determinant is

$$\Delta = \theta_1[\{1-\exp(\theta_2 \xi_{12})\}\xi_{11}\exp(\theta_2 \xi_{11}) - \{1-\exp(\theta_2 \xi_{11})\}\xi_{12}\exp(\theta_2 \xi_{12})], \tag{26}$$

the modulus of which has a maximum when

$$\xi_{12} = \xi_1(\max), \tag{27}$$

$$\xi_{11} = -\frac{1}{\theta_2^*} - \frac{\xi_{12}\exp(\theta_2^* \xi_{12})}{1-\exp(\theta_2^* \xi_{12})}. \tag{28}$$

Substituting these values into equation (25) and letting γ_2 be the fraction by which the yield at time $\xi_{12} = \xi_1(\max)$ falls short of the maximum obtainable, we find that the optimum value of ξ_{11} corresponds to a yield falling short of the maximum by a fraction

$$\gamma_1 = \exp\{-[1+\gamma_2 \log \gamma_2/(1-\gamma_2)]\}. \tag{29}$$

In particular if the second observation is taken after a sufficiently long time so that the asymptotic value is almost achieved, then γ_2 is close to zero and $\gamma_1 = 1/e$, so that the first observation should be made when the process is 63·2 % complete. As a second example, if the second observation is taken at a time when the process is 50 % complete, then $\gamma_1 = 0{\cdot}73$ and the first observation should be made at a time when the process is 27 % complete.

3·2. *Mitscherlich equation*

Consider the model

$$\eta = \theta_1 + \theta_2 \exp(\theta_3 \xi_1) \tag{30}$$

for $$0 \leqslant \xi_1(\min) \leqslant \xi_{11} < \xi_{12} < \xi_{13} \leqslant \xi_1(\max).$$

With $\theta_3 < 0$ and $\theta_2 < 0$ this expression represents the well-known Mitcherlich law of diminishing returns. For example, in crop experiments with ξ_1 the amount of added fertilizer and η the amount of growth, $\theta_1 + \theta_2$ represents the growth when no fertilizer is added and θ_1 the hypothetical growth obtained from an infinite amount of fertilizer; θ_3 measures the rate at which the response to each additional increment of fertilizer decreases. The same model has been employed to relate the weight η of chickens to their cumulative feed consumption ξ_1 (Hendricks *et al.* 1932). With $\theta_3 > 0$ this model has been used to represent early growth (Brant, 1951).

The value of the determinant is

$$\begin{aligned}\Delta = \theta_2[&\xi_{11} \exp(\theta_3 \xi_{11}) \{\exp(\theta_3 \xi_{13}) - \exp(\theta_3 \xi_{12})\} + \xi_{12} \exp(\theta_3 \xi_{12}) \\ &\times \{\exp(\theta_3 \xi_{11}) - \exp(\theta_3 \xi_{13})\} + \xi_{13} \exp(\theta_3 \xi_{13}) \{\exp(\theta_3 \xi_{12}) - \exp(\theta_3 \xi_{11})\}].\end{aligned} \tag{31}$$

It is readily shown that irrespective of the signs of θ_1, θ_2, θ_3, the optimum design is obtained with

$$\xi_{11} = \xi_1(\min), \tag{32}$$

$$\xi_{12} = -\frac{1}{\theta_3^*} + \frac{\xi_{11} \exp(\theta_3^* \xi_{11}) - \xi_{13} \exp(\theta_3^* \xi_{13})}{\exp(\theta_3^* \xi_{11}) - \exp(\theta_3^* \xi_{13})}, \tag{33}$$

$$\xi_{13} = \xi_1(\max). \tag{34}$$

In particular in a fertilizer experiment if $\xi_{11} = \xi_1(\min)$ corresponded to no dressing of fertilizer and $\xi_{13} = \xi_1(\max)$ corresponded to a dressing for which the response fell short of the maximum by a certain fraction γ_2, then ξ_{12} would be a dressing for which the response fell short of the maximum possible by a fraction γ_1 where again γ_1 would be given by (29). When $\xi_1(\max)$ is sufficiently large to give an almost maximal response then the intermediate value of the design should be that expected to yield a response of 63·2 % of the maximum possible.

3·3. *Simple first-order decay with rate a function of temperature*

Consider the model

$$\eta = \exp(-\theta_1 \xi_1 e^{-\theta_2 \xi_2}) \tag{35}$$

for $$0 \leqslant \xi_1(\min) \leqslant \xi_{11} < \xi_{12} \leqslant \xi_1(\max),$$
$$\xi_2(\min) \leqslant \xi_{21} < \xi_{22} \leqslant \xi_2(\max).$$

This would apply for example where η represented the fractional amount remaining at time $t = \xi_1$ of a material decaying according to the first-order law, when the rate was temperature-

sensitive. Assuming that the temperature effect follows the Arrhenius rule, θ_1 is the specific rate at some absolute temperature T_0, θ_2 is proportional to the 'activation energy' and $\xi_2 = T^{-1} - T_0^{-1}$ is the deviation of the reciprocal of the temperature, T, from the reciprocal of the base temperature T_0.

We find for preliminary values θ_1^*, θ_2^* that

$$\phi_1(\xi_1, \xi_2) = -\xi_1 \exp\{-(\theta_2^* \xi_2 + \theta_1^* \xi_1 e^{-\theta_2^* \xi_2})\} = \frac{\eta^* \ln \eta^*}{\theta_1^*},$$

$$\phi_2(\xi_1, \xi_2) = -\theta_1^* \xi_2 \phi_1(\xi_1, \xi_2) = -\xi_2 \eta^* \ln \eta^*, \tag{36}$$

where $\eta_u^* = \exp(-\theta_1^* \xi_{1u} e^{-\theta_2^* \xi_{2u}})$, the expected fraction of decaying substance remaining in the uth experiment. Thus

$$\text{mod}\,\Delta = \text{mod} \begin{vmatrix} \frac{\eta_1^*}{\theta_1^*} \ln \eta_1^* & -\xi_{21} \eta_1^* \ln \eta_1^* \\ \frac{\eta_2^*}{\theta_1^*} \ln \eta_2^* & -\xi_{22} \eta_2^* \ln \eta_2^* \end{vmatrix} = \frac{\eta_1^* \ln \eta_1^* \eta_2^* \ln \eta_2^*}{\theta_1^*} (\xi_{22} - \xi_{21}). \tag{37}$$

Now $-\eta \log \eta$ has its maximum value when $\eta = 1/e$. Also, if the range of permissible values for the times, ξ_1, is sufficiently wide then the expression (37) is maximized by ensuring that $\xi_{22} - \xi_{21}$ shall achieve the maximum possible range $\xi_2(\text{max}) - \xi_2(\text{min})$ and adjusting the times of reaction so that the anticipated fractions remaining, η_1^*, η_2^*, are each equal to $1/e$. The actual values of ξ_{11} and ξ_{12} required in this case are obtained by substituting $\eta = e^{-1}$ in equation (35). The appropriate times ξ_{11}, ξ_{12}, associated with $\xi_2(\text{min})$ and $\xi_2(\text{max})$ are then given by

$$\xi_{11} = \frac{1}{\theta_1^*} \exp\{\theta_2^* \xi_2(\text{min})\},$$

$$\xi_{12} = \frac{1}{\theta_1^*} \exp\{\theta_2^* \xi_2(\text{max})\}. \tag{38}$$

It is instructive to consider a numerical example. Suppose the absolute temperature is restricted to the range 380–420° A. It is then convenient to take as an arbitrary base temperature the value $T_0 = 400°$ A. Suppose with time measured in minutes a guessed value for the rate θ_1 at this arbitrary base temperature is $\theta_1^* = 1{\cdot}0$ and a guessed value for $\theta_2^* = 16{,}000$, then we have

$$\theta_1^* = 1{\cdot}0, \quad \theta_2^* = 16{,}000$$

$$380° \text{ A.} \leqslant T = (\xi_2 + 0{\cdot}0025)^{-1} \leqslant 420° \text{ A.}$$

The optimum design will consist of two experiments, the first of which is performed at the minimum temperature of 380° A for time $\xi_{11} = 8{\cdot}21$ min. and the second at the maximum temperature of 420° A for 0·15 min. These times are found by substituting in equation (38) and are simply those values at which, if the guessed values are correct and there were no experimental error, the fraction of product remaining would be $1/e$. The design is illustrated geometrically in Fig. 4. In this figure the design locus in the plane of $\phi_1(\xi_1, \xi_2)$, $\phi_2(\xi_1, \xi_2)$ is bounded by the values $\phi_1(\xi_1, \xi_2) = 0$ and $\phi_1(\xi_1, \xi_2) = -1/\theta_1^* e$, the largest and the smallest values, respectively, that this derivative can take. The straight lines in Fig. 4 are for constant ξ_2 (i.e. for constant temperature) and the curved lines for constant ξ_1 (i.e. for constant time). Where, as we have assumed above, the permissible range of ξ_2 dictates the

design points selected, the design locus is restricted to that portion of the plane within the triangle bounded by the three straight lines

$$\phi_1(\xi_1, \xi_2) = -1/\theta_1^* e,$$
$$\phi_2(\xi_1, \xi_2) = -\theta_1^* \phi_1(\xi_1, \xi_2)\, \xi_2(\min),$$
$$\phi_2(\xi_1, \xi_2) = -\theta_1^* \phi_1(\xi_1, \xi_2)\, \xi_2(\max).$$

The triangle so delineated is itself the triangle of largest area which can be formed with two points on the design locus and the origin with only temperature restricted.

The appropriate design for the case considered is indicated by the two heavy dots in the diagram. It will be seen from the figure that, if ξ_1 were the restricting variable, then a somewhat different result would be obtained, since now the triangle of maximum area would be achieved when the value of $\phi_1(\xi_1, \xi_2)$ fell somewhat short of $-1/\theta_1^* e$.

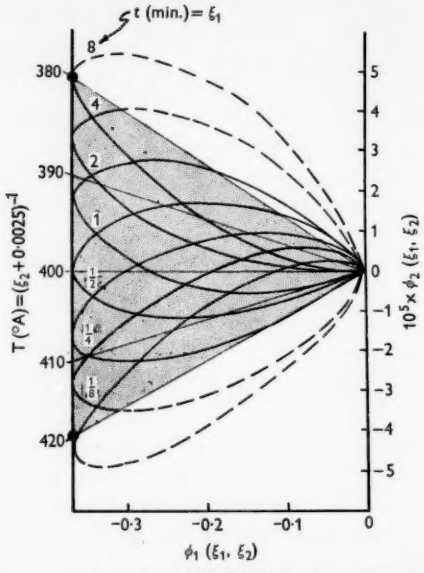


Fig. 4. The design locus for $\eta = \exp\{-1{\cdot}0te^{-16{,}000(T^{-1}-0{\cdot}0025)}\}$. The optimal solution for $T(\min) = 380°$A., $T(\max) = 420°$ A. and $0 \leqslant \xi_1 < \infty$ is indicated by heavy dots.

4. Calculation of the adjustments

Suppose a design of this type has been obtained for assumed values of the parameters $\boldsymbol{\theta}^*$, and suppose that the N calculated values at these sets of conditions for $\boldsymbol{\theta} = \boldsymbol{\theta}^*$ are given by the column vector $\boldsymbol{\eta}^*$. Suppose finally that observed responses from the N experiments performed are provided by the column vector $\mathbf{y}$. Then proceeding in the usual way a $p \times 1$ vector of adjustments $\boldsymbol{\delta}$ to be applied to the preliminary values $\boldsymbol{\theta}^*$ are given to a first approximation by

$$\boldsymbol{\delta} = (\mathbf{F}^{*\prime}\mathbf{F}^*)^{-1}\mathbf{F}^{*\prime}(\mathbf{y} - \boldsymbol{\eta}^*)$$

or in the special case where the number of observations N is equal to the number of parameters p by

$$\boldsymbol{\delta} = \mathbf{F}^{*-1}(\mathbf{y} - \boldsymbol{\eta}^*).$$

In the particular example of § 3·3,

$$\mathbf{F}^* = \frac{1}{e}\begin{bmatrix} -1/\theta_1^* & \xi_{21} \\ -1/\theta_1^* & \xi_{22} \end{bmatrix}$$

and

$$\mathbf{y} - \boldsymbol{\eta}^* = \begin{bmatrix} y_2 - e^{-1} \\ y_1 - e^{-1} \end{bmatrix},$$

so that the adjusted values will be

$$\theta_1' = \theta_1^* + \delta_1,$$
$$\theta_2' = \theta_2^* + \delta_2,$$

where δ_1 and δ_2 are the 'regression coefficients' of the elements in $\mathbf{y} - \boldsymbol{\eta}^*$ on the columns of the matrix $\mathbf{F}^*$.

To see the meaning of this suppose that the value of θ_1^* were taken at a base temperature such that $\xi_{21} = -\xi_{22}$, then the adjustments in θ_1^* and θ_2^* are simply

$$\delta_1 = -(\bar{y} - e^{-1})\, e\theta_1^*,$$
$$\delta_2 = (y_1 - y_2)\, e/2\xi_{21}.$$

As would be expected, therefore, when the observed values of y were equal but different from the value e^{-1}, this would imply that adjustment was needed only in the basic rate constant θ_1^*. Unequal values for y_1 and y_2 on the other hand would imply the necessity for adjustment in θ_2^*.

5. Discussion

When parameters are jointly estimated from observations made at particular sets of conditions, it frequently happens that the estimates are highly correlated and consequently have very large variances. This may be simply the result of poor design in which case the first group of estimates may be used as starting values to design a more informative set of trials from which greatly improved estimates can be obtained. Not infrequently, however, it turns out that the unsatisfactory nature of the estimates is a property of the model itself, and that even the best design still yields highly correlated estimates. This point is well illustrated by models such as

$$\eta = \frac{1 + \theta_1 \xi_1}{\theta_2 + \theta_3 \xi_1}.$$

In accordance with the method already outlined, the optimal design will be found by choosing three points on the design locus which together with the origin form the tetrahedron of greatest volume. But the design locus for such a model often turns out to be a nearly straight line, so that for any reasonable range of the variable, ξ_1, the volume of the tetrahedron and hence the value of the determinant is small. This arises from the nature of the model itself. Whatever values of the ξ's are chosen in such examples, when one has found one set of values of the θ's that most closely agrees with a given set of data, one can, by making suitable compensating changes in the parameters, obtain other quite different values of the θ's that agree almost as well.

In this connexion it is important to distinguish between the problem we are considering here, which is that of finding designs most efficient for estimating the parameters $\boldsymbol{\theta}$, and a somewhat different problem which is *not* discussed here, that of finding designs such that the response $\eta = f(\boldsymbol{\xi}; \boldsymbol{\theta})$ is determined most precisely in a given region. A particular design may be such that the estimates, $\boldsymbol{\theta}$, of the parameters are highly correlated with large

variances and yet the response, η, is estimated reasonably accurately. The reason for this is not difficult to see. Instability in the $\boldsymbol{\theta}$'s which is *compensating* is by its very nature not such as will cause instability in the estimate of η.

The authors wish to acknowledge their indebtedness to a referee and to E. M. L. Beale for valuable suggestions for improving the presentation.

REFERENCES

Box, G. E. P. (1957). Use of statistical methods in the elucidation of basic mechanisms. *Bull. Inter. Statist. Inst.* **36**, part 3, 215–25.

Brant, J. W. A. (1951). Rate of early growth in domestic fowl. *Poult. Sci.* **30**, 343–61.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Ann. Math. Statist.* **24**, 586–602.

Elfving, G. (1952). Optimum allocation in linear regression theory. *Ann. Math. Statist.* **23**, 255–62.

Hendricks, W. A., Jull, M. A. & Titus, H. W. (1932). The utilization of feed by chickens. I. The requirements for growth. *Poult. Sci.* **2**, 74–7.

Koshal, R. S. (1933). Application of the method of maximum likelihood to the improvement of curves fitted by the method of least squares. *J. Roy. Statist. Soc.* **96**, 303–13.

Wald, A. (1943). On the efficient design of statistical investigations. *Ann. Math. Statist.* **14**, 134–40.

THE ESTIMATION OF MISSING AND MIXED-UP OBSERVATIONS IN SEVERAL EXPERIMENTAL DESIGNS

By J. D. BIGGERS

Department of Physiology, Royal Veterinary College, London

I. Introduction

The simple analysis of most experimental designs depends upon their fully balanced arrangement. Often observations are lost for a few of the units in the course of an experiment, and to overcome the consequent difficulties in the statistical analysis, because of non-orthogonality, the technique of inserting 'missing values' was introduced (Allan & Wishart, 1930; Yates, 1933). The values which are inserted minimize the error sum of squares.

In this paper the problem of calculating these missing values will be investigated from a general point of view for the following designs:

(1) randomized block,
(2) randomized block design with replication within units,
(3) cross-over,
(4) latin square and higher squares,
(5) split-plot.

Adjustment of the variances of treatment comparisons is sometimes necessary after estimating missing values, but although this can be done by an extension of the present procedure it will not be discussed in this paper.

Notation

Scalars—letters in italics.

Vectors—lower case letters in Clarendon type.

Matrices—capital letters in Clarendon type.

Transposition of matrices and vectors is indicated by a prime.

Summation—two summation symbols will be used. If we consider the sequence $x_i (i = 1, 2, \dots n)$, we define ordinary summation over all values as $\Sigma_i x_i$. Also we define

$$\Sigma_{(i)} x_i = x_1 + x_3 + x_{n-1}, \quad \text{say},$$

to mean summation over some subset of the subscript values—typically one relating to the missing values.

Associated observations—an observation in a design is denoted by $X_{ijk\dots}$, where the number of subscripts indicates the order of classification involved. Thus an observation in a randomized block requires two subscripts for its specification, and an observation in a Graeco-latin square requires four.

If we consider two observations they may have certain subscripts in common. This property enables us to describe the relation between pairs of observations; observations related in a specific way will be called *associates*. For example, we consider two observations X_{fgh} and X_{uvw}. The following classes of associate can be recognized:

(1) zero-order associates: $f \neq u, g \neq v, h \neq w$;

(2) 1st-order associates: (i) i associates $f = u, g \neq v, h \neq w$;
(ii) j associates $f \neq u, g = v, h \neq w$;
(iii) k associates $f \neq u, g \neq v, h = w$;

(3) 2nd-order associates: (i) ij associates $f = u, g = v, h \neq w$;
(ii) ik associates $f = u, g \neq v, h = w$;
(iii) jk associates $f \neq u, g = v, h = w$.

Particular associates may or may not exist depending on the experimental design.

A *potential observation* in a design will be indicated by a capital letter: $X_{ij...}$; if this is in fact a *missing observation* it will be indicated in lower case: $x_{ij...}$.

General solution

Let there be p independent missing values denoted by $x_i (i = 1, 2, ..., p)$. The error sum of squares is a quadratic function of these values, $F(x_1, x_2, ..., x_p)$. The vector of partial derivatives of this function is given by

$$2(\mathbf{x}'\mathbf{A} - \mathbf{q}')/N, \tag{1}$$

where $\mathbf{x}$ is a column vector of the p missing values,

$\mathbf{A}$ is a $p \times p$ symmetric matrix whose elements are determined by the experimental design and the distribution of the missing values,

$\mathbf{q}$ is a column vector whose elements are calculated from the available observations,

N is a constant determined by the experimental design.

The values of the x_i where $F(\mathbf{x})$ is a minimum are given by the solution of p simultaneous equations obtained by equating each of its partial derivatives to zero. These are, in matrix notation:

$$\mathbf{A}\mathbf{x} = \mathbf{q}. \tag{2}$$

Several methods may be used to solve equation (2). Most textbooks on the design and analysis of experiments give the formula for one missing observation in randomized block, cross-over and latin-square designs, and the special formulae for two or three missing observations have also been published for the randomized block design (Baten, 1939, 1952; Federer, 1951, 1955) and the latin-square design (Federer, 1955). The split-plot design has been dealt with for single missing observations by Anderson (1946) and Khargonkar (1948). Usually, where more than one missing observation exists in a design, it is recommended that the estimates be obtained by an iterative process based on the formula for one missing observation. Other methods are the covariance method (Bartlett, 1937; Coons, 1957) and the method of diagonal iteration (Healy & Westmacott, 1956) for use with automatic computors. None of these methods entails the explicit determination of the simultaneous equations. In this paper rules will be given which enable (2) to be written down rapidly for the different experimental designs. The main advantage of this procedure is that it enables the nature of the equations to be examined and this leads to simplifications in several important complex cases.

The matrix method has been the subject of a paper by Thompson (1956) on missing values in the randomized block design. This worker, however, has only dealt with a restricted distribution of missing values, presumably because values of $\mathbf{A}^{-1}$ are written down instead of $\mathbf{A}$. Very recently Wilkinson (1957) has shown how the matrix $\mathbf{A}$ can be used to obtain exact variances of treatment comparisons, and has also extended the method for the estimation of missing observations in the analysis of covariance.

II. RANDOMIZED BLOCK DESIGN

Suppose we have r blocks each containing t plots to which t treatments are allotted at random. We denote the observation made on the jth treatment in the ith block by $X_{ij} (i = 1, 2, \ldots, r;\ j = 1, 2, \ldots, t)$. At the end of the experiment let p of these values be missing. It should be noted that the missing observations are indicated by only some values of the subscripts.

Let the ith block and the jth treatment contain c_i and d_j missing observations respectively. Then

$$p = \Sigma_i c_i = \Sigma_j d_j.$$

Missing observations in the same block are i or *block associates* and missing observations receiving the same treatment are called j or *treatment associates*. The missing observation x_{gh} will have therefore $(c_g - 1)$ block associates, $(d_h - 1)$ treatment associates and $\{p - (c_g + d_h - 1)\}$ zero associates.

Let B_i = total of the available observations in the ith block,
T_j = total of the available observations receiving the jth treatment,
G = total of all available observations.

The error term in the analysis of variance is given by

$$\Sigma_{(i)}\Sigma_{(j)} x_{ij}^2 - \Sigma_{(i)} (B_i + \Sigma_{(j)} x_{ij})^2/t - \Sigma_{(j)} (T_j + \Sigma_{(i)} x_{ij})^2/r + (G + \Sigma_{(i)}\Sigma_{(j)} x_{ij})^2/rt + \text{constant}. \quad (3)$$

Partial differentiation of (3) with respect to a particular missing observation, x_{gh}, and equating the derivative to zero, gives the equation

$$rtx_{gh} - r\Sigma_{(j)} x_{gj} - t\Sigma_{(i)} x_{ih} + \Sigma_{(i)}\Sigma_{(j)} x_{ij} = rB_g + tT_h - G.$$

Separation of terms corresponding to zero, i and j associates gives

$$(r-1)(t-1)x_{gh} + \overbrace{(1-r)\sum_{j \neq h} x_{gj}}^{i \text{ associates}} + \overbrace{(1-t)\sum_{i \neq g} x_{ih}}^{j \text{ associates}} + \overbrace{\sum_{i \neq g}\sum_{j \neq h} x_{ij}}^{\text{zero associates}} = rB_g + tT_h - G. \quad (4)$$

Expression (4) enables us to write down the coefficients of the x_{ij}, the B_g and T_h, and hence the matrix **A** and the vector **q** in (2). The rules which determine the elements of **A** are:

(i) $(r-1)(t-1)$ for the missing observation under consideration,
(ii) $(1-r)$ for the i or block associates,
(iii) $(1-t)$ for the j or treatment associates,
(iv) 1 for the zero associates.

In all practical situations the matrix **A** will be non-singular, and thus the equation (2) will have a unique solution.

Example 1. Table 1 shows the results of a randomized block experiment ($r = 8$; $t = 5$). No missing values occurred in this experiment, but four have been omitted to provide the example. The missing values are calculated as follows:

(1) Write out the table of results so that the rows correspond to separate blocks. This is essential to avoid confusion in determining the associates.

(2) Obtain the coefficients of the associates. In the example these are 28 for the missing observation under consideration, -7 for the i associates, -4 for the j associates and 1 for the zero associates.

(3) Insert in the table of results the letter x_{ij} for the missing value with subscripts determined by the position of the missing value in the table, e.g. x_{35} is the missing value in

the third block receiving the fifth treatment. Space should be left to insert the estimated missing value.

Table 1. *The wet weight in mg* (x) *of embryonic chick tibiae after cultivation on a chemically defined medium containing five different concentrations of glucose. Treatment numbers are placed in parentheses. The variate analysed is log* $10x$

Block	Glucose concentration (mg./ml.)					B_i	Completed block totals
	0·5 (1)	1·0 (2)	2·0 (3)	4·0 (4)	8·0 (5)		
I	0·88	1·15	1·33	1·39	1·57	—	6·32
II	1·06	1·22	1·63	1·54	1·37	—	6·82
III	0·97	1·34	(x_{33} = 1·50)	1·66	(x_{35} = 1·53)	3·97 (B_3)	7·00
IV	1·09	1·21	1·16	1·50	1·48	—	6·44
V	1·14	1·37	1·58	1·52	1·44	—	7·05
VI	1·13	1·33	1·65	1·57	1·47	—	7·15
VII	1·00	1·21	(x_{73} = 1·44)	1·45	1·52	5·18 (B_7)	6·62
VIII	1·12	1·30	1·35	(x_{84} = 1·56)	1·61	5·38 (B_8)	6·94
T_j	—	—	8·70 (T_3)	10·63 (T_4)	10·46 (T_5)	48·31 (G)	—
Completed treatment totals	8·39	10·13	11·64	12·19	11·99	—	54·34

Table 2. *Calculation of* **A** *and* **q**

Subscripts	**A**				**q**		
	33	35	73	84	B_i	T_j	G
33	28	−7	−4	1	3·97	8·70	48·31
35	−7	28	1	1	3·97	10·46	48·31
73	−4	1	28	1	5·18	8·70	48·31
84	1	1	1	28	5·38	10·63	48·31
					r = 8	t = 5	−1

(4) Obtain the B_i, T_j and G from the available observations and write them at the border of the table.

(5) Prepare a table as in Table 2. List the subscripts in a row at the top and a column at the left-hand side of the table. By comparing the subscripts in pairs determine which subscripts are common. This determines the type of association, and the appropriate coefficient is written down at the intersection of the corresponding row and column. At the right of the

table columns for the B_i, T_j and G are prepared. The i subscript in the left-hand column determines the B_i total in a given row, while the j subscript determines the T_j total. Beneath the B_i and T_j columns the values of r and t are written down respectively, and beneath the G column the value -1 is written.

(6) Invert the matrix **A**. If the exact variances of treatment comparisons are not required $\mathbf{A}^{-1}$ need not be obtained explicitly, and it is quicker to solve the equations (2) directly with **q** determined as in step (7) below (see, for example, Fox (1950)).

(7) Write down the solution as follows:

$$\begin{matrix} & & \mathbf{A^{-1}} & & & \\ \end{matrix}$$

$$\begin{bmatrix} x_{33} \\ x_{35} \\ x_{73} \\ x_{84} \end{bmatrix} = \begin{bmatrix} 0{\cdot}039 & 0{\cdot}010 & 0{\cdot}005 & -0{\cdot}002 \\ 0{\cdot}010 & 0{\cdot}038 & 0{\cdot}000 & -0{\cdot}002 \\ 0{\cdot}005 & 0{\cdot}000 & 0{\cdot}037 & -0{\cdot}001 \\ -0{\cdot}002 & -0{\cdot}002 & -0{\cdot}001 & 0{\cdot}036 \end{bmatrix} \begin{matrix} \mathbf{q} \\ \begin{bmatrix} 26{\cdot}95 \\ 35{\cdot}75 \\ 36{\cdot}63 \\ 47{\cdot}88 \end{bmatrix} \end{matrix}$$

q is obtained by row by row multiplication of each row of the right-hand side of Table 2 with the bottom row. Thus, in the example, the upper entry of **q** is given by

$$8 \times 3{\cdot}97 + 5 \times 8{\cdot}70 - 48{\cdot}31 = 26{\cdot}95.$$

The values of the x_{ij} are then obtained by standard row by column multiplication. In the example $x_{33} = 1{\cdot}50$; $x_{35} = 1{\cdot}53$; $x_{73} = 1{\cdot}44$; $x_{84} = 1{\cdot}56$.

(8) The values of the x_{ij} are inserted in Table 1, and the block and treatment totals completed. The standard analysis of variance then follows. In this example 4 D.F. are subtracted from the error term.

Special cases. The process of calculating $\mathbf{A}^{-1}$ can be simplified if **A** is a $p \times p$ matrix of the form:

$$\begin{bmatrix} a & b & . & . & . & b \\ b & a & . & . & . & b \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ b & b & . & . & . & a \end{bmatrix}. \tag{5}$$

This is a circulant matrix and hence the sum of the elements of a row is a factor of its determinant. Thus if $\{a+b(p-1)\} = 0$ the matrix is singular. This property will be used later.

The inverse of (5) is given by

$$\frac{1}{(a-b)\{a+b(p-1)\}} \begin{bmatrix} \{a+b(p-2)\} & -b & . & . & . & -b \\ -b & \{a+b(p-2)\} & . & . & . & -b \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ -b & -b & . & . & . & \{a+b(p-2)\} \end{bmatrix}$$

Thus once **A** is known, $\mathbf{A}^{-1}$ can be written down directly.

The special cases where this is useful in the randomized block design are:

(1) all missing observations are zero associates,

(2) all missing observations are either block or treatment associates. These matrices also occur in other experimental designs.

III. Randomized block design with replication within units

Suppose we have r blocks each containing t plots to which t treatments are allotted at random, and that in each plot there are s independent observations. The analysis of variance takes the form shown in Table 3. At the present time there is considerable discussion on the linear models which are assumed in the analysis of this design (Anderson & Bancroft, 1952; Wilk & Kempthorne, 1955). The models determine the expectations of the mean squares in the analysis of Table 3, and these in turn determine the appropriate error terms in the tests of significance. There are three error terms to consider:

(1) the within-plots sum of squares (E_I) (Case 1),

(2) the blocks × treatments sum of squares (E_{II}) (Case 2),

(3) ($E_I + E_{II}$) (Case 3).

Table 3. *Analysis of variarce for a $r \times t$ randomized block design with s observations per plot*

Source of variation	D.F.	Sum of squares
Between blocks	$(r-1)$	$(B-C)$
Between treatments	$(t-1)$	$(T-C)$
Blocks × treatments interaction (E_{II})	$(r-1)(t-1)$	$(P-B-T+C)$
Within plots (E_I)	$rt(s-1)$	$(S-P)$
Total	$(rst-1)$	$(S-C)$

Where S = total sum of squares, C = correction for the mean, B = block sum of squares, T = treatment sum of squares, P = plot sum of squares.

We denote the kth observation on the jth treatment in the ith block by

$$X_{ijk} (i = 1, 2, \ldots, r; j = 1, 2, \ldots, t; k = 1, 2, \ldots, s).$$

At the end of the experiment let p_{ij} observations be missing in each plot. Then the total number of missing observations is $p = \Sigma_i \Sigma_j p_{ij}$.

The following associates only exist: zero, i, j and ij.

Let P_{ij} = total of the available observations in the ijth plot,

B_i = total of the available observations in the ith block,

T_j = total of the available observations receiving the jth treatment,

G = total of all available observations.

The three sets of equations for estimating the missing values when each of the error terms are minimized have been obtained by an analysis similar to that applied to the randomized block design. The rules which determine the elements of **A** and **q** for each case are shown in Table 4.

If we write down (2) for these cases, ordering the elements of **A** in groups corresponding to the ij, we obtain in general:

$$\begin{bmatrix} \mathbf{A}_{11}^{11} & \mathbf{A}_{12} & . & . & \mathbf{A}_{1.rt} \\ \mathbf{A}_{12} & \mathbf{A}_{22}^{12} & . & . & \mathbf{A}_{2.rt} \\ . & . & . & . & . \\ . & . & . & . & . \\ \mathbf{A}_{1.rt} & \mathbf{A}_{2.rt} & . & . & \mathbf{A}_{rt.rt}^{rt} \end{bmatrix} \begin{bmatrix} \mathbf{x}^{11} \\ \mathbf{x}^{12} \\ . \\ . \\ \mathbf{x}^{rt} \end{bmatrix} = \begin{bmatrix} \mathbf{q}^{11} \\ \mathbf{q}^{12} \\ . \\ . \\ \mathbf{q}^{rt} \end{bmatrix}, \tag{6}$$

where $\mathbf{A}_{\alpha\beta}^{ij}(\alpha = \beta)$ are submatrices along the main diagonal, one corresponding to each plot. Each is a square symmetric matrix of order p_{ij}. The superscript denotes the plot, while the subscript denotes the position of the submatrix,

$\mathbf{A}_{\alpha\beta}(\alpha \neq \beta)$ are rectangular submatrices,

$\mathbf{x}^{ij}$ is a vector of the p_{ij} missing values in plot ij,

$\mathbf{q}^{ij}$ is a vector of p_{ij} elements calculated from the available observations.

Table 4. *Coefficients of associates and typical elements of* $\mathbf{q}$ *for the three cases in the randomized block design with replication within units*

Associate	Coefficient		
	Case 1 (E_I)	Case 2 (E_{II})	Case 3 ($E_I + E_{II}$)
Missing value	$(s-1)$	$(r-1)(t-1)$	$(rst-r-t+1)$
ij	-1	$(r-1)(t-1)$	$(1-r-t)$
i	0	$(1-r)$	$(1-r)$
j	0	$(1-t)$	$(1-t)$
Zero	0	1	1
	Typical elements of $\mathbf{q}$		
Case 1	P_{fg}		
Case 2	$rB_f + tT_g - rtP_{fg} - G$		
Case 3	$rB_f + tT_g - G$		

If a plot has no missing values the submatrices of the corresponding row and column in (6) vanish. The elements of the submatrices along the main diagonal are determined by the coefficients for the missing value and its ij associates only.

We now consider the three cases discussed above.

Case 1. The matrix of (6) is of the form

$$\begin{bmatrix} \mathbf{A}_{11} & . & . & . & . \\ . & \mathbf{A}_{12} & . & \mathbf{O} & . \\ . & . & . & . & . \\ . & \mathbf{O} & . & . & . \\ . & . & . & . & \mathbf{A}_{rt.rt} \end{bmatrix} \tag{7}$$

Since the inverse of (7) is

$$\begin{bmatrix} \mathbf{A}_{11}^{-1} & . & . & . & . \\ . & \mathbf{A}_{12}^{-1} & . & \mathbf{O} & . \\ . & . & . & . & . \\ . & \mathbf{O} & . & . & . \\ . & . & . & . & \mathbf{A}_{rt.rt}^{-1} \end{bmatrix},$$

the estimates of the missing values in plot ij is independent of the missing values in all other plots. Therefore the missing values for the ijth plot are given by the solution of

$$\underset{p_{ij}\times p_{ij}}{\begin{bmatrix} (s-1) & -1 & . & . & -1 \\ -1 & (s-1) & . & . & -1 \\ . & . & . & . & . \\ . & . & . & . & . \\ -1 & -1 & . & . & (s-1) \end{bmatrix}} \begin{bmatrix} x_{ij1} \\ x_{ij2} \\ . \\ . \\ x_{ijp_{ij}} \end{bmatrix} = \begin{bmatrix} P_{ij} \\ P_{ij} \\ . \\ . \\ P_{ij} \end{bmatrix}.$$

This has the solution $P_{ij}/(s-p_{ij})$ for all the x_{ijk}. Thus in this case the missing observations are estimated by the average of the available observations in the plot. If all the values in a plot are missing $p_{ij}=s$ and the equations have no straightforward solution.

When $r=1$ we have the case of a fully randomized design. Thus an estimate of a missing value, if required, is the average of the results available on the particular treatment.

Case 2. The submatrices, in this case, have equal elements throughout. The $\mathbf{A}^{ij}_{\alpha\beta}(\alpha=\beta)$ have elements equal to $(r-1)(t-1)$ while the $\mathbf{A}_{\alpha\beta}(\alpha\neq\beta)$ have elements equal to $(1-r)$, $(1-t)$ or 1. The rank of matrix $\mathbf{A}$ in all practical cases is rt since it consists of rt sets of identical rows of elements. Thus to solve (6) we may allot within each plot $(p_{ij}-1)$ arbitrary values and then solve for the remainder. In practice, therefore, we would insert arbitrary values for all but one missing value in each plot, say $P_{ij}/(s-p_{ij})$ for $s\neq p_{ij}$, and then write out (6) as though one value only is missing in each plot. Alternatively, if we work entirely with plot totals we can use the formula for a randomized block design to estimate a missing plot, ignoring all the available observations on the plot. This situation arises when we minimize the whole-plot error term in a split-plot design.

Case 3. Here the matrix of (6) is non-singular and cannot be simplified. In this case the matrix remains non-singular even if all observations are missing from a plot, thus allowing estimates of these values to be made.

IV. Cross-over design

Consider an $r\times t$ design, where each of the t columns is a separate replicate of the r treatments, subject to the condition that each treatment occurs an equal number of times (a) in each row. Then $t=ar$.

Denote each observation by $X_{ijk}(i,k=1,2,\ldots,r;\ j=1,2,\ldots,t)$. Let p observations be missing. In this design the following associates only exist: zero, i, j, k, ik.

Let R_i = total of the available observations in the ith row,
C_j = total of available observations on the jth column,
T_k = total of the available observations receiving the kth treatment,
G = total of all available observations.

The following rules determine the elements of $\mathbf{A}$:

(i) $(r-1)(t-2)$ for the missing observation under consideration,
(ii) $2(1-r)$ for the ik associates,
(iii) $(2-r)$ for the i and k associates,
(iv) $(2-t)$ for the j associates,
(v) 2 for the zero associates.

A typical element of $\mathbf{q}$ is given by

$$rR_f+tC_g+rT_h-2G.$$

When $r = 2$ the coefficient for i and k associates is zero. This leads to independence of the equations for estimating the missing values in some instances, thus allowing the formula for a single missing value to be used separately for each plot. A similar situation arises in split-plot designs (see Example 2).

V. LATIN-SQUARE DESIGN

The formulae for a $t \times t$ latin-square design can be obtained from those for the cross-over design by putting $a = 1$. Under this restriction ik associates do not exist. Thus the rules which determine the elements of **A** for the latin-square design are:

(i) $(t-1)(t-2)$ for the missing observation under consideration,

(ii) $(2-t)$ for the i, j and k associates,

(iii) 2 for the zero associates.

A typical element of **q** is then $t(R_f + C_g + T_h) - 2G$.

VI. HIGHER ORDER SQUARE DESIGNS

Suppose we have a $t \times t$ square of order n, where n denotes the number of classifications. Denote the observation made on each plot by $X_{\alpha_1 \alpha_2 \dots \alpha_n} (\alpha_i = 1, 2, \dots, t;\ i = 1, 2, \dots, n)$. A particular observation will be denoted by $X_{\beta_1 \beta_2 \dots \beta_n}$, where the β_i represent fixed values of the α_i. Let p observations be missing.

In all these designs only zero and 1st order associates exist.

Let C_{α_i} = total of the available observations on the α_ith classification,

G = total of all available observations.

By an extension of the methods already described it can be shown that a typical equation of (2) is

$$\begin{aligned}(t-1)(t+1-n)\, x_{\beta_1 \beta_2 \dots \beta_n} &+ (n-1-t) \sum_{\alpha_2 + \beta_2} \sum_{\alpha_3 + \beta_3} \dots \sum_{\alpha_n + \beta_n} x_{\beta_1 \alpha_2 \dots \alpha_n} \\ &+ (n-1-t) \sum_{\alpha_1 + \beta_1} \sum_{\alpha_3 + \beta_3} \dots \sum_{\alpha_n + \beta_n} x_{\alpha_1 \beta_2 \dots \alpha_n} \\ &+ \quad \dots \quad \dots \quad \dots \\ &+ (n-1-t) \sum_{\alpha_1 + \beta_1} \sum_{\alpha_2 + \beta_2} \dots \sum_{\alpha_{n-1} + \beta_{n-1}} x_{\alpha_1 \alpha_2 \dots \beta_n} \\ &+ (n-1) \sum_{\alpha_1 + \beta_1} \sum_{\alpha_2 + \beta_2} \dots \sum_{\alpha_n + \beta_n} x_{\alpha_1 \alpha_2 \dots \alpha_n} \\ &= tC_{\beta_1} + tC_{\beta_2} + \dots + tC_{\beta_n} - (n-1)G. \end{aligned} \tag{8}$$

From (8) we may determine the rules shown in Table 5 for writing down the coefficients of matrix **A** for various values of n. $n = 2$ determines a randomized block design where the number of treatments equals the number of blocks. $n = 3$ determines the coefficients of a $t \times t$ latin square, thus giving an alternative proof of the formula given earlier.

If in the higher square designs $(t-1)$ 1st order associates only are missing then matrix **A** is of the form (5). The sum of the elements in a row is

$$(t-1)(t-n+1) + (t-1)(n-1-t) = 0,$$

and therefore the matrix **A** is singular and equation (2) has no straightforward solution. This situation arises whenever all the missing values are in one classification, e.g. a complete row, column or treatment in a latin square. A full description of methods for dealing with missing values in latin-square designs where **A** is singular is given by De Lury (1946).

Table 5. *Coefficients of associates for square designs of different order*

No. of constraints	2	3	4	5	n
Associate	Randomized block $s = r = t$	Latin square	Graeco-latin square	Hyper-graeco-latin square	—
Missing observation	$(t-1)(t-1)$	$(t-1)(t-2)$	$(t-1)(t-3)$	$(t-1)(t-4)$	$(t-1)(t-n+1)$
1st order	$(1-t)$	$(2-t)$	$(3-t)$	$(4-t)$	$(n-1-t)$
Zero	1	2	3	4	$(n-1)$

VII. Split-plot designs

(1) *Whole-plot treatments arranged in randomized blocks*

Consider a design in which r blocks containing t plots to which t whole-plot treatments are allotted at random. Further let s split-plot treatments be allotted at random within each whole plot. The analysis of variance of this design is given in Table 6. There are two error terms to be considered: (1) E_I corresponding to the whole-plots (Case 1), and (2) E_{II} corresponding to the split-plots (Case 2).

Table 6. *Analysis of variance for a split-plot design with whole-plot treatments arranged in randomized blocks*

Source of variation	D.F.	Sum of squares
Between blocks	$(r-1)$	$(B-C)$
Between whole-plot treatments (A)	$(t-1)$	$(S-C)$
Error (E_I)	$(r-1)(t-1)$	$(W-B-S+C)$
Total	$(rt-1)$	$(W-C)$
Between split-plot treatments (B)	$(s-1)$	$(T-C)$
AB interaction	$(s-1)(t-1)$	$(I-S-T+C)$
Error (E_{II})	$t(s-1)(r-1)$	$(Z-W-I+S)$
Total	$(rst-1)$	$(Z-C)$

Where Z = total sum of squares of split-plots, W = total sum of squares of whole plots, C = correction for the mean, B = block sum of squares, S = whole-plot treatment sum of squares, T = split-plot treatment sum of squares, I = treatment combination sum of squares.

We denote the observation made on the kth split-plot treatment in the ijth plot by X_{ijk} ($i = 1, 2, \ldots, r$; $j = 1, 2, \ldots, t$; $k = 1, 2, \ldots, s$). At the end of the experiment let p split-plots be missing.

The following associates exist: zero, i, j, k, ij, ik, jk.

Case 1. Table 6 shows that the expression for the error term is the same as that in Case 2 of the generalized randomized block design. The solution in the two cases is identical. We ignore the available observations and estimate new whole-plot values using the formula for a randomized block design.

Case 2. Let W_{ij} = total of the available observations in the ijth plot,

I_{jk} = total of the available observations receiving the jth whole-plot treatment and the kth split-plot treatment,

S_j = total of the available observations receiving the jth whole-plot treatment.

The rules which determine the elements of **A** are:

(i) $(r-1)(s-1)$ for the missing observation under consideration,
(ii) $(1-s)$ for the jk associates,
(iii) $(1-r)$ for the ij associates,
(iv) 1 for the j associates,
(v) 0 for the zero i, k and ik associates.

A typical element of **q** is given by:

$$rW_{fg}+sI_{gh}-S_g.$$

(2) *Whole-plot treatments in a $t \times t$ latin square*

Consider a $t \times t$ latin square to which whole-plot treatments are allotted at random, and further let s split-plot treatments be allotted within each whole plot. The analysis of variance is shown in Table 7. As before there are two error terms to be considered.

Table 7. *Analysis of variance for a split-plot design with whole-plot treatments arranged in a latin square*

Source of variation	D.F.	Sum of squares
Between rows	$(t-1)$	$(R-C)$
Between columns	$(t-1)$	$(K-C)$
Between whole-plot treatments (A)	$(t-1)$	$(S-C)$
Error (E_I)	$(t-1)(t-2)$	$(W-R-K-S+2C)$
Total	(t^2-1)	$(W-C)$
Between split-plot treatments (B)	$(s-1)$	$(T-C)$
AB interaction	$(t-1)(s-1)$	$(I-S-T+C)$
Error (E_{II})	$t(t-1)(s-1)$	$(Z-W-I+S)$
Total	(st^2-1)	$(Z-C)$

Where Z = total sum of squares of split-plots, W = total sum of squares of whole plots, C = correction for the mean, R = row sum of squares, K = column sum of squares, S = whole-plot treatment sum of squares, T = split-plot treatment sum of squares, I = treatment combination sum of squares.

We denote an observation on each split-plot by $X_{ijkl}\,(i,j,k=1,2,\ldots,t;\ l=1,2,\ldots,s)$, where i denotes the row, j the column, k the whole-plot treatment associated with the plot ij, and l the split-plot treatment. Let p split-plots be missing.

The following associates only exist due to the restrictions on a latin-square design: zero, $i, j, k, l, il, jl, kl, ijk$.

Case 1. If we wish to minimize the whole-plot error we ignore the available observations and estimate new whole-plot values using the formula for a latin-square design.

Case 2. Let W_{ij} = total of the available observations on the ijth plot,

I_{kl} = total of the available observations receiving the kth whole-plot treatment and the lth split-plot treatment,

S_k = total of the available observations receiving the kth whole-plot treatment.

The rules which determine the elements of **A** are:

(i) $(t-1)(s-1)$ for the missing observation under consideration,

(ii) $(1-t)$ for the ijk associates,

(iii) $(1-s)$ for the kl associates,

(iv) 1 for the k associates,

(v) 0 for the zero, i, j, l, il and jl associates.

A typical element of **q** is given by $tW_{ef}+sI_{gh}-S_g$.

The fact that the coefficient of some associates is zero in these split-plot designs allows the problem of solving the equations to be simplified in many cases. If the matrix **A** is written down ordering the elements in groups according to the j subscript for the randomized block design and the k subscript for the latin-square design a matrix of the form (7) is obtained. This consists of submatrices along the main diagonal, one corresponding to each j or k, and zeros elsewhere. As these submatrices may be separately inverted to give the required solution considerable reduction in the labour of computation results. Further, in many simple cases the missing observations may be fitted quite independently using the standard formula for one missing observation. The independence of missing values within whole plots in these designs has been pointed out previously by Pearce (1953). If an entire whole plot is missing the $t \times t$ submatrix will have $(r-1)(t-1)$ along the main diagonal and $(1-r)$ elsewhere. The sum of a row is zero, and thus the submatrix is singular and there is no straightforward solution to the equations.

Example 2. Table 8 shows the results of a split-plot experiment where the whole-plot treatments are arranged in randomized blocks ($r=4$; $t=5$; $s=2$). Missing values have been artificially chosen to provide the example. Table 9 is prepared from the data of Table 8 in exactly the same way as described in Example 1, except that the subscripts have been ordered using the middle (j) subscript. From the matrix **A** it is seen that the estimation of x_{121} and x_{222} is independent of the estimation of x_{132} and x_{232}. Thus instead of inverting a 4×4 matrix we invert two 2×2 matrices, considerably simplifying the problem.

The estimates are given by

$$\begin{bmatrix} x_{121} \\ x_{222} \end{bmatrix} = 1/8 \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 4{\cdot}74 \\ 4{\cdot}74 \end{bmatrix}$$

and

$$\begin{bmatrix} x_{132} \\ x_{232} \end{bmatrix} = 1/8 \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2{\cdot}56 \\ 3{\cdot}76 \end{bmatrix}.$$

Thus $x_{121} = 1{\cdot}19$; $x_{222} = 1{\cdot}19$; $x_{132} = 1{\cdot}43$; $x_{232} = 1{\cdot}73$.

Table 8. *The wet weight in mg (x) of embryonic chick tibiae after cultivation on a chemically defined medium containing different concentrations of glucose or mannose. Treatment numbers are placed in parentheses. The variate analysed is log 10x*

Block	Split-plot treatment	Whole-plot treatment Hexose concentration (mg./ml.)				
		0·5 (1)	1·0 (2)	2·0 (3)	4·0 (4)	8·0 (5)
I	Glucose (1)	0·88	(x_{121} = 1·19)	1·33	1·39	1·57
	Mannose (2)	0·78	1·15	(x_{132} = 1·43)	1·42	1·51
II	Glucose (1)	1·06	1·22	1·63	1·54	1·37
	Mannose (2)	1·09	(x_{222} = 1·19)	(x_{232} = 1·73)	1·47	1·44
III	Glucose (1)	0·97	1·34	1·44	1·66	1·59
	Mannose (2)	0·90	1·16	1·44	1·48	1·59
IV	Glucose (1)	1·09	1·21	1·16	1·50	1·48
	Mannose (2)	1·07	1·32	1·36	1·41	1·41

Table 9. *Calculation of* **A** *and* **q**

Subscripts	A				q		
	121	222	132	232	W_{ij}	I_{jk}	S_j
121	3	1	0	0	1·15	3·77	7·40
222	1	3	0	0	1·22	3·63	7·40
132	0	0	3	−1	1·33	2·80	8·36
232	0	0	−1	3	1·63	2·80	8·36
					$r = 4$	$s = 2$	−1

VIII. MIXED-UP OBSERVATIONS

Sometimes the identity of observations may be lost so that we only know their total. These are known as mixed-up values. Bose & Mahalanobis (1938) and Nair (1940) have discussed the analysis of randomized block and latin-square designs where some observations are mixed up, and they have shown that estimates may be obtained which minimize the error mean square.

In the general case both missing and mixed-up values may occur. We may relate this to the previous sections by supposing that some of the p missing observations are not missing entirely, but that certain of them have been mixed up. Then the p missing values may be connected by c relations of the form:

$$\Sigma_{(p)} b_{ij} x_i = u_j \quad (i = 1, 2, \ldots, p; j = 1, 2, \ldots, c) \tag{9}$$

The coefficients of (9) are either 0 or 1. For example, if $p = 5$, and the missing values are x_2, x_7, x_9, x_{13} and x_{19}, with x_7 and x_9 and also x_2 and x_{19} mixed up and x_{13} entirely missing, we have the relations:

$$0.x_2+1.x_7+1.x_9+0.x_{13}+0.x_{19} = u_1$$

and

$$1.x_2+0.x_7+0.x_9+0.x_{13}+1.x_{19} = u_2.$$

To obtain estimates for the missing and mixed-up values we minimize the error mean square subject to the conditional equations (9). Although the coefficients of (9) are restricted to 0 and 1 to deal with the case of mixed-up values, this restriction is not essential in the following discussion. The problem will be treated generally in matrix notation.

Equation (8) may be written:

$$\mathbf{Bx} = \mathbf{u}, \tag{10}$$

where $\mathbf{B}$ is a $p \times c$ matrix of coefficients; each column of this matrix will contain one unit element and $(c-1)$ zero elements, $\mathbf{x}$ is a column vector of the p values to be estimated, $\mathbf{u}$ is a column vector of c mixed-up totals.

If we introduce a column vector $(2\boldsymbol{\lambda}/N)$ of c undetermined multipliers the Lagrange equations are given by

$$\frac{\partial}{\partial \mathbf{x}}\{F(x_1, x_2, \ldots, x_p)\} + 2\boldsymbol{\lambda}'(\mathbf{Bx}-\mathbf{u}) = \mathbf{O}.$$

Using the notation of equation (1) we find that

$$2/N(\mathbf{x}'\mathbf{A}-\mathbf{q}') + 2/N\boldsymbol{\lambda}'\mathbf{B} = \mathbf{O},$$

and after transposition and rearrangement

$$\mathbf{Ax}+\mathbf{B}'\boldsymbol{\lambda} = \mathbf{q}. \tag{11}$$

The values of the p missing values which minimize the error mean square are given by the simultaneous solution of (10) and (11). These two matrix equations may be combined to give

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{B}' \\ \hline \mathbf{B} & \mathbf{O} \end{array}\right] \left[\begin{array}{c} \mathbf{x} \\ \hline \boldsymbol{\lambda} \end{array}\right] = \left[\begin{array}{c} \mathbf{q} \\ \hline \mathbf{u} \end{array}\right],$$

the solution of which is given by:

$$\left[\begin{array}{c} \mathbf{x} \\ \hline \boldsymbol{\lambda} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{B}' \\ \hline \mathbf{B} & \mathbf{O} \end{array}\right]^{-1} \left[\begin{array}{c} \mathbf{q} \\ \hline \mathbf{u} \end{array}\right].$$

Example 3. Suppose that in Example 2, x_{121} and x_{132} have been confused and that the combined weights after transformation to logs is 2·50. Similarly, suppose that x_{222} and x_{232} are also confused, and that the combined weight is 2·83.

The matrix of coefficients is found to be

$$\begin{array}{c} \\ \mathbf{A} \\ \\ \\ \\ \mathbf{B} \\ \\ \end{array} \left[\begin{array}{cccc|cc} 3 & 1 & 0 & 0 & 1 & 0 \\ 1 & 3 & 0 & 0 & 0 & 1 \\ 0 & 0 & 3 & -1 & 1 & 0 \\ 0 & 0 & -1 & 3 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{array}\right] \begin{array}{c} \\ \\ \\ \mathbf{B}' \\ \\ \mathbf{O} \end{array}$$

Then the missing observations are given by

$$\begin{bmatrix} x_{121} \\ x_{222} \\ x_{132} \\ x_{232} \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \tfrac{1}{6} \begin{bmatrix} 1 & 0 & -1 & 0 & 3 & -1 \\ 0 & 1 & 0 & -1 & -1 & 3 \\ -1 & 0 & 1 & 0 & 3 & 1 \\ 0 & -1 & 0 & 1 & 1 & 3 \\ 3 & -1 & 3 & 1 & -8 & 0 \\ -1 & 3 & 1 & 3 & 0 & -8 \end{bmatrix} \begin{bmatrix} 4{\cdot}74 \\ 4{\cdot}74 \\ 2{\cdot}56 \\ 3{\cdot}76 \\ 2{\cdot}50 \\ 2{\cdot}83 \end{bmatrix}.$$

Thus $x_{121} = 1{\cdot}14$; $x_{222} = 1{\cdot}16$; $x_{132} = 1{\cdot}36$; $x_{232} = 1{\cdot}67$.

I wish to thank Dr D. R. Cox for helpful comments on the draft of this paper.*

REFERENCES

ALLAN, F. E. & WISHART, J. (1930). *J. Agric. Sci.* **20**, 399.

ANDERSON, R. L. (1946). *Biometrics*, **2**, 41.

ANDERSON, R. L. & BANCROFT, T. A. (1952). *Statistical Theory in Research*. London: McGraw-Hill.

BARTLETT, M. S. (1937). *J.R. Statist. Soc.* Suppl. **4**, 137.

BATEN, W. D. (1939). *Technical Bulletin* 165, Michigan State College, Agricultural Experiment Station.

BATEN, W. D. (1952). *Biometrics*, **8**, 42.

BOSE, S. S. & MAHALANOBIS, P. C. (1938). *Sankhyā*, **4**, 103.

COONS, I. (1957). *Biometrics*, **13**, 387.

DE LURY, D. B. (1946). *J. Amer. Statist. Ass.* **41**, 370.

FEDERER, W. T. (1951). *Research Bulletin* 380, Iowa State College, Agricultural Experiment Station.

FEDERER, W. T. (1955). *Experimental Design*. New York: Macmillan and Co.

FOX, L. (1950). *J.R. Statist. Soc.* B, **12**, 120.

FOX, L. & HAYES, J. G. (1951). *J.R. Statist. Soc.* B, **13**, 83

HEALY, M. & WESTMACOTT, M. (1956). *Appl. Statist.* **5**, 203.

KHARGONKAR, S. A. (1948). *J. Indian Soc. Agric. Statist.* **1**, 147.

NAIR, K. R. (1940). *Sankhyā*, **4**, 581.

PEARCE, S. C. (1953). *Field Experimentation with Fruit Trees and other Perennial Plants*. Commonwealth Agricultural Bureau.

THOMPSON, H. R. (1956). *Biometrics*, **12**, 241.

WILK, M. B. & KEMPTHORNE, O. (1955). *J. Amer. Statist. Ass.* **50**, 1144.

WILKINSON, G. N. (1957). *Biometrics*, **13**, 363.

WILKINSON, G. N. (1958*a*). *Biometrics*, **14**, 257.

WILKINSON, G. N. (1958*b*). *Biometrics*, **14**, 360.

YATES, F. (1933). *Emp. J. Exp. Agric.* **1**, 129.

* Since this paper was submitted for publication two further papers have appeared by Wilkinson (1958*a*, *b*) in which the problem of estimating missing values is discussed more generally. In these papers Wilkinson also deals with missing values in balanced and partially balanced non-orthogonal designs, and discusses the analysis of those cases where **A** is singular.

NOTE ON A THREE-DECISION TEST FOR COMPARING TWO BINOMIAL POPULATIONS

By J. H. DARWIN

Department of Scientific and Industrial Research, Wellington, New Zealand

1. Introduction and summary

Suppose that in each of two processes there is a probability of an event happening and it is desired to compare these probabilities. Pairs of trials of these processes may be made in sequence and a test made on the results. There may be four possible outcomes of each pair of trials—the event may happen in both processes, in neither, or in only one. If the result of the test is to be that just one of the processes will be used in the future the first two outcomes are irrelevant. If the result of the test is to be that just one of the processes will be used in the future if it is markedly superior to the other, but that either may be used if they do not appear to differ importantly then all outcomes are relevant.

In this paper we produce a test for this latter case. Another test of this kind has recently been described by Armitage (1957). He does not employ pairs of trials having the same outcome and consequently has a different form of test. Both his and our tests are truncated in the sense that a maximum number of pairs of trials is needed for each.

When the probability of the same result for each process is well known, our test will require a smaller maximum number of pairs of trials for the same power, and the same probability of rejecting when true the null hypothesis that the processes are equally effective. A numerical example of this is given.

When this probability of the same result for each process is not known Armitage's test is preferable since it does not depend on a knowledge of this probability. When, however, the tests are being used to compare a control process with a new process our test which, in its construction assumes some value of the probability of the same result, will be affected only favourably should the control value of this probability depart from the assumed value.

The strongest connexion between our test and Armitage's is that the boundaries which determine when our test stops are a special case of the boundaries required by his test. We use exact matrix theory to find the probabilities of the boundaries being crossed, and he uses approximate normal theory. For some comparative examples we calculated, the probability given by this approximate theory differed from those given by the exact theory by at most 0·0015.

We also give an approximate set of power contours which enable tests of our type to be constructed for any value of the power and of the probability of rejecting the null hypothesis when it is true.

2. Derivation of the test

Suppose the probabilities of the events happening are p_1 and p_2 for the two processes.

Then in a pair of trials the probability of the event happening in the first process but not in the second is $p_1(1-p_2) = \pi_1$, say.

The probability of the same result in the two trials is $p_1 p_2 + (1-p_1)(1-p_2) = \pi_2$, say.

The probability of the event happening in the second process but not in the first is $p_2(1-p_1) = \pi_3$, say.

We consider a variable γ taking values 1, 0 and -1 for these three cases, respectively. Suppose m pairs of trials have been made in sequence and $\gamma(m)$ is the sum of the m values of γ. We could imagine $\gamma(m)$ drawn on a graph as a function of m. Suppose $\gamma(0) = \frac{1}{2}(n+1)$ where n is an odd integer, and suppose the probability that $\gamma(m)$ has the value Y is $P_{Y,m}$. Then

$$P_{Y,m+1} = \pi_1 P_{Y-1,m} + \pi_2 P_{Y,m} + \pi_3 P_{Y+1,m}. \tag{1}$$

If $\gamma(m)$ crosses $Y = 0$ we shall suppose the first process is accepted as that to use in the future. If $\gamma(m)$ crosses $Y = n+1$ we shall suppose the second process is accepted. If $\gamma(m)$ has crossed neither of these boundaries when M pairs of trials have been made the processes will be regarded as equally acceptable for future use and the sequence stopped. We now wish to find values of M and n for which there is a high probability of a markedly superior process being accepted, and a high probability of either process being permitted in the future if there is the same probability of the event happening in each process. To find these values we solve the system (1) with $P_{1-a,m} = P_{n+b,m} = 0$ for all integers a and b greater than 0; i.e. for $\gamma(m)$ confined between 1 and n. The $P_{Y,m}$ satisfy the matrix equation

$$\mathbf{P}_{m+1} = \begin{pmatrix} \pi_2 & \pi_3 & . & . & . & . \\ \pi_1 & \pi_2 & \pi_3 & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & \pi_1 & \pi_2 & \pi_3 \\ . & . & . & . & \pi_1 & \pi_2 \end{pmatrix} \mathbf{P}_m$$

$$= D_n \mathbf{P}_m, \quad \text{say},$$

where $\mathbf{P}_m$ is the vector of the $P_{Y,m}$.

Suppose $|D_n - \lambda I|$ is called $D_n(\lambda)$. Then $D_u(\lambda)$ satisfies the difference equation

$$D_u(\lambda) = (\pi_2 - \lambda) D_{u-1}(\lambda) - \pi_1 \pi_3 D_{u-2}(\lambda)$$

with solution

$$D_u(\lambda) = \frac{(\frac{1}{2}[\pi_2 - \lambda + \{(\pi_2-\lambda)^2 - 4\pi_1\pi_3\}^{\frac{1}{2}}])^{u+1} - (\frac{1}{2}[\pi_2 - \lambda - \{(\pi_2-\lambda)^2 - 4\pi_1\pi_3\}^{\frac{1}{2}}])^{u+1}}{\{(\pi_2-\lambda)^2 - 4\pi_1\pi_3\}^{\frac{1}{2}}}.$$

It easily follows that the latent roots of D_n are

$$\lambda_r = \pi_2 + 2\sqrt{(\pi_1\pi_3)} \cos\frac{r\pi}{n+1} \quad (r = 1, 2, \ldots, n). \tag{3}$$

It then follows that $D_u(\lambda_r) = (\pi_1\pi_3)^{\frac{1}{2}u} \sin[(u+1)\, r\pi/(n+1)]$.

The elements of the latent column vector C_r corresponding to λ_r satisfy the equation

$$\pi_1 v_j + (\pi_2 - \lambda) v_{j+1} + \pi_3 v_{j+2} = 0.$$

It then follows that the jth element of C_r is proportional to $(-)^j D_{j-1}(\lambda_r)/\pi_3^j$; and, similarly, the jth element of the latent row vector R_r corresponding to λ_r is proportional to $(-)^j D_{j-1}(\lambda_r)/\pi_1^j$.

A spectral resolution of the matrix D_n is then possible (see e.g. Bartlett 1955, p. 26) and we may write

$$D_n^m = \sum_{r=1}^{n} \lambda_r^m \frac{C_r R_r'}{C_r' R_r}. \tag{4}$$

The probability of $\gamma(m)$ crossing $Y = 0$ up to the time when M pairs of trials have been made, if $\gamma(m)$ has previously been between 1 and n, is

$$\sum_{m=0}^{M-1} P_{1m}\pi_3 = \pi_3 \sum_m \sum_{r=1}^{n} \lambda_r^m C_{r,1} R_{r,\frac{1}{2}(n+1)}/(C_r' R_r)$$
$$= \frac{2}{n+1}\pi_3\left(\frac{\pi_3}{\pi_1}\right)^{\frac{1}{2}(n-1)} \sum_{r=1}^{n} \frac{1-\lambda_r^M}{1-\lambda_r} \sin\frac{r\pi}{n+1}\sin\frac{r\pi}{2}. \tag{5}$$

Similarly, the probability of $\gamma(m)$ crossing the boundary $Y = n+1$ up to the time when M pairs of trials have been made, if $\gamma(m)$ has previously been between 1 and n, is

$$\frac{2}{n+1}\pi_1\left(\frac{\pi_1}{\pi_3}\right)^{\frac{1}{2}(n-1)} \sum_{r=1}^{n} \frac{1-\lambda_r^M}{1-\lambda_r} \sin\frac{r\pi}{n+1}\sin\frac{r\pi}{2}. \tag{6}$$

The probability of $\gamma(m)$ crossing the line $m = M$ between $Y = 1$ and $Y = n$ without having previously crossed either $Y = 0$ or $Y = n+1$ is

$$\frac{2}{n+1}\left(\pi_3\left(\frac{\pi_3}{\pi_1}\right)^{\frac{1}{2}(n-1)} + \pi_1\left(\frac{\pi_1}{\pi_3}\right)^{\frac{1}{2}(n-1)}\right) \sum_{r=1}^{n} \frac{\lambda_r^M}{1-\lambda_r} \sin\frac{r\pi}{n+1}\sin\frac{r\pi}{2}. \tag{7}$$

3. Calculations and discussion

Values of M and n given in Table 1 have been found to satisfy these requirements:

(i) For some π_2 and for a value of π_1/π_3 greater than 1, the probability (6) was made equal to 0·95 or 0·80. (8)

(ii) For the same π_2 and for $\pi_1 = \pi_3$ the probability (7) was made equal to 0·95. (8′)

That is, the probability of accepting, when true, the null hypothesis of no difference between the processes, is 0·95, and the power of the test when π_1 and π_3 differ is either 0·95 or

Table 1. *Distance between acceptance limits* (n) *and number of trials* (M) *which will ensure the power indicated, at the* 5 % *significance level*

	Power	$\pi_1/\pi_3 (= S) = 1{\cdot}2$	1·4	1·6	1·8	2·0	2·5	3·0
$\pi_2 = 0{\cdot}25$	0·95	n 181	99	71	57	49	37	31
		M 2200	656	340	2200	162	95	68
	0·80	n 141	77	55	45 (43)	37	29	25
		M 1336	399	206	136 (131)	96	59	43
$\pi_2 = 0{\cdot}50$	0·95	n 181	99	71	57	49	37	31
		M 3301	986	511	331	244	144	103
	0·80	n 141	77	55	45	37	29	25
		M 2006	600	311	205	145	89	65
$\pi_2 = 0{\cdot}75$	0·95	n 181	99	71	57	49	37	31
		M 6606	1972	1025	666	491	290	209
	0·80	n 141	77	55	45	37	29	25
		M 4018	1201	623	412	292	179	132

0·80. The conditions (8) determine unique values of M and n. The range of π_1/π_3 considered was 1 to 3.

That value of n was chosen for which the difference between the values of M for which (i) and (ii) of (8) were satisfied was a minimum. Then for the sake of definiteness that value of M was taken for which the power was closest to 0·95 or 0·80. The two values of n given at one place under $\pi_1/\pi_3 = 1{\cdot}8$ indicate that $n = 43$, $M = 131$ is the better choice, but $n = 45$, $M = 136$ has been put in for comparison with the lower figures in this column.

The calculations were not as laborious as might be imagined, since in all cases at most six and usually three terms of the series in (7) had to be taken, and always for $\pi_1 > \pi_3$, (5) was negligible.

In (6)

$$\frac{2}{n+1}\pi_1\left(\frac{\pi_1}{\pi_3}\right)^{\frac{1}{2}(n-1)}\sum_r \frac{\sin(r\pi/(n+1))\sin\frac{1}{2}r\pi}{1-\lambda_r} = \frac{\pi_1^{\frac{1}{2}(n+1)}}{\pi_1^{\frac{1}{2}(n+1)}+\pi_3^{\frac{1}{2}(n+1)}}, \tag{9}$$

since this is the known probability of $\gamma(m)$ crossing $Y = n+1$ when M is infinite. We used this as a partial check on the sufficiency of the numbers of figures we were keeping.

There are good approximations available to the probabilities (6) and (7). First, there is the normal approximation. In this $\gamma(M)$ for large M is taken as normally distributed with mean $M(\pi_1-\pi_3)$ and variance $M(1-\pi_2-(\pi_1-\pi_3)^2)$. The probability of $\gamma(m)$ crossing the boundary $Y = n+1$ before M pairs of trials have been made is found (after consideration of the paths that reach $m = M$ with Y less than $n+1$ but which have previously crossed $Y = n+1$) to be

$$1-F\left[\frac{\frac{1}{2}(n+1)-(\pi_1-\pi_3)M}{\{M(1-\pi_2-(\pi_1-\pi_3)^2)\}^{\frac{1}{2}}}\right]+\exp\left[\frac{(n+1)(\pi_1-\pi_3)}{1-\pi_2-(\pi_1-\pi_3)^2}\right]F\left[\frac{-\frac{1}{2}(n+1)-(\pi_1-\pi_3)M}{\{M(1-\pi_2-(\pi_1-\pi_3)^2)\}^{\frac{1}{2}}}\right], \tag{10}$$

where

$$F(x) = \int_{-\infty}^{x} e^{-\frac{1}{2}x^2}\,dx/\sqrt{(2\pi)}. \tag{11}$$

This result is quoted by Armitage (1957) from work by Bartlett (1946). We compared the powers it gave for the values of n and M given by the first part of the line for $\pi_2 = 0{\cdot}75$ in Table 1 with the powers given by the exact quantity (6) and in none of the seven cases was the difference as much as 0·0015.

Another approach to an approximation is given by considering (7) directly. Because of (9) we may write (7) in the form

$$\frac{\frac{1}{n+1}\sum_r \frac{(1-\pi_2)\lambda_r^M}{1-\lambda_r}\sin\frac{r\pi}{n+1}\sin\frac{r\pi}{2}}{\frac{1}{n+1}\sum_r \frac{(1-\pi_2)}{1-\lambda_r}\sin\frac{r\pi}{n+1}\sin\frac{r\pi}{2}}. \tag{12}$$

Now suppose for convenience we denote π_1/π_3 by S. Then,

$$1-\lambda_r = (1-\pi_2)\left(1-\frac{2\sqrt{S}}{1+S}\cos\frac{r\pi}{n+1}\right). \tag{13}$$

For small r and large n we may write this as

$$\begin{aligned}&\frac{1-\pi_2}{(n+1)^2}\left(\left(1-\frac{2\sqrt{S}}{1+S}\right)(n+1)^2+\frac{2\sqrt{S}}{1+S}\frac{r^2\pi^2}{2}\right)\\ &\quad=\frac{1-\pi_2}{(n+1)^2}\left(\left(1-\frac{2\sqrt{S}}{1+S}\right)(n+1)^2+\frac{r^2\pi^2}{2}\right)\left(1-\frac{(1-2\sqrt{S}/(1+S))\frac{1}{2}r^2\pi^2}{(1-2\sqrt{S}/(1+S))(n+1)^2+\frac{1}{2}r^2\pi^2}\right).\end{aligned} \tag{14}$$

For the range of S considered, 3 to 1·2, $1-2\sqrt{S}/(1+S)$ goes from 0·8660 to 0·9959. For the values of n and r encountered this meant that the last factor of (14) departs from one by only a small amount. We only had to consider values of r less than 7 so that r had a range virtually independent of n. This means we may write the denominator of (12) approximately as

$$\sum_{r=1}^{6} \frac{r\pi \sin(\frac{1}{2}r\pi)}{(n+1)^2(1-2\sqrt{S}/(1+S)+\frac{1}{2}r^2\pi^2)}. \tag{15}$$

This exhibits the denominator as a function of $(n+1)^2(1-2\sqrt{S}/(1+S))$ only. We shall use this approximation below. Again, $M\log\lambda_r$ for low r is approximately

$$M\log(\pi_2+2\sqrt{(\pi_1\pi_3)}) - \frac{r^2\pi^2 2M\sqrt{(\pi_1\pi_3)}}{2(n+2)^2(\pi_2+2\sqrt{(\pi_1\pi_3)})} \tag{16}$$

$$\simeq -M(1-\pi_2)(1-2\sqrt{S}/(1+S)) - \tfrac{1}{2}M(1-\pi_2)^2(1-2\sqrt{S}/(1+S))^2 - \frac{r^2\pi^2M(1-\pi_2)}{2(n+1)^2}\frac{2\sqrt{S}}{1+S}(1+(1-\pi_2)(1-2\sqrt{S}/(1+S))). \tag{17}$$

The two leading terms of (17) each have M and π_2 occurring only in the combination $M(1-\pi_2)$. The remaining terms are proportional to $1-2\sqrt{S}/(1+S)$. In particular if $S=1$, $1-2\sqrt{S}/(1+S)=0$, and the approximation given by the leading terms is especially good. We can in fact write approximately for (7)

$$\frac{4}{\pi}\left(e^{-f}-\frac{e^{-9f}}{3}+\frac{e^{-25f}}{5}-\ldots\right), \tag{18}$$

where $f=\pi^2M(1-\pi_2)/(n+1)^2$. The value of f making (18) equal to 0·95 is 0·24557.

Suppose now we have found M and n for the test to have a given power for a high value of π_2, say 0·75, and for S not equal to one. Then $M\log\lambda_r(0{\cdot}75)$ for $\pi_2=0{\cdot}75$ will be larger than $M_1\log\lambda_r(0{\cdot}25)$ for $\pi_2=0{\cdot}25$ and $M_1=\frac{1}{3}M$ because both the second-order approximations in (17) have a negative sign. However $(M_1-1)\log\lambda_r(0{\cdot}25)-M_1\log\lambda_r(0{\cdot}25)$ is approximately $0{\cdot}75(1-2\sqrt{S}/(1+S))+r^2\pi^2 0{\cdot}75\sqrt{S}/\{(1+S)(n+1)^2\}$ and this in our examples is of the same order as the second terms in (17). In fact a drop of only one or two from M_1 was needed to balance the negative terms. This explains why in Table 1, for fixed S and n we are able to achieve the same power when $M(1-\pi_2)$ is kept approximately constant.

We may write the term

$$\frac{r^2\pi^2M(1-\pi_2)}{2(n+1)^2}\frac{2\sqrt{S}}{1+S}$$

of (17) as

$$\frac{r^2\pi^2M(1-\pi_2)}{2(n+1)^2}\left(1-\left(1-\frac{2\sqrt{S}}{1+S}\right)\right). \tag{19}$$

The term in $1-2\sqrt{S}/(1+S)$ of (19) is of the same order as the second-order term in (17). It now appears from (15), (17) and (19) that the power (the complement of (7)) is approximately a function of the two quantities

$$\frac{M(1-\pi_2)}{(n+1)^2} \quad \text{and} \quad (n+1)^2(1-2\sqrt{S}/(1+S)). \tag{20}$$

The second-order term in (17) and those neglected in (15) offset each other to some extent. If they are neglected we may, following (20), solve approximately for the M and n required by (8), the equations

$$\frac{M(1-\pi_2)}{(n+1)^2} = \text{constant},$$

$$(n+1)^2(1-2\sqrt{S}/(1+S)) = \text{constant}.$$

For the values of n and S from Table 1 for a power of 0·95 and for increasing n, $(n+1)^2(1-2\sqrt{S}/(1+S))$ takes values 137·19, 139·33, 142·98, 140·23, 139·93, 139·87 and 137·17. For a power of 0·80 the values are 90·57, 86·86, 82·58, 88·21, 84·65, 85·10 and 83·50.

For $\pi_2 = 0{\cdot}75$, a power of 0·95 and for increasing n, $M(1-\pi_2)/(n+1)^2$ takes values 0·0510, 0·0502, 0·0491, 0·0495, 0·0494, 0·0493 and 0·0499. For a power of 0·80 the values are 0·0488, 0·0497, 0·0506, 0·0487, 0·0497, 0·0494 and 0·0498.

The variations in these quantities are caused partly by the influence of the term neglected in arriving at (20), but mainly by the fact that M and n have to be integers and n has to be odd. This means that the powers for the values of n and M given in the table are never exactly those stated. They diverge from them by up to four or five in the third place of decimals for low n and less for large n.

We would not expect (20) to hold if S were to increase much beyond 3, since $2\sqrt{S}/(1+S)$ is starting to diverge too far from one for $1-2\sqrt{S}/(1+S)$ to be considered small compared with one.

4. Comparison with Armitage's test

Armitage's test has a given power for any fixed π_1/π_3, regardless of the value of π_2. Our test has a given power for a fixed π_1/π_3 and a given value of π_2. If our test is designed for a given value of π_2 and this turns out to be the correct value it can be expected that our test will require a smaller maximum number of observations for the same power. We give a numerical example of this.

In his Table 5 is given the maximum number of pairs of trials with unequal results to give a power of 0·95 and a probability 0·05 of rejection of a true null hypothesis. To make the comparison with our test we divide this maximum number by $1-\pi_2$ to take account of the discarded pairs of trials in which the processes gave equal results. We may divide this number by our value of M corresponding to this value of π_2. The ratio then is roughly constant at 1·27. This represents, for this power and probability of rejection of a true null hypothesis, the relative total number of trials required by the two tests when the assumed π_2 is equal to the actual π_2.

Suppose, however, the true value of π_2 differs from that assumed in the construction of our test. Then if the true value is bigger than that assumed the power will decrease, if it is smaller it will increase. We can make an estimate of how big the change in power is likely to be when π_2 varies, by recalling the approximate result (20). If this holds we can take any values of M and n and map the power function of the test by varying π_2 and S. We have done this and present the result in Fig. 1. From this figure we can get approximate values of M and n for a test in which the power takes any value between 0·50 and 0·99 and the probability of acceptance of a true null hypothesis lies between 0·917 and 0·993.

Fig. 1 shows that there may be a marked drop in power if π_2 is much bigger than was assumed. This is a serious weakness of the test if there is no knowledge of the likely value. Then it will be safer to use Armitage's test.

However, we consider one case in which the lack of knowledge of π_2 may prove no bar to the use of our test. Suppose we are comparing a standard process of fairly well-known properties with an untried process. Suppose for the control process the probability of the event happening is P_1 and for the new process the probability is P_2. By reversing if necessary the definition of the happening of the event determined by a process we may suppose P_1 is greater than one-half. An improvement on the control process may correspond to a value of P_2 either (*a*) greater than P_1; or (*b*) less than P_1.

We suppose (*a*) holds and that it is decided that an important improvement over control will correspond to values of P_2 greater than P_2', say. The argument that follows will hold, mutatis mutandis, if (*b*) holds. The quantities P_1 and P_2' are enough to determine π_2 and $\pi_1/\pi_3 (= S)$ and so to give values M and n corresponding to any power and any level of rejection of a true null hypothesis.

Consider what happens if P_2 is different from P_2'. If it decreases below P_2',

$$\pi_2 \equiv P_1 P_2 + (1-P_1)(1-P_2)$$

will decrease because P_1 is greater than one half, and $\pi_1/\pi_3 \equiv P_2(1-P_1)/\{P_1(1-P_2)\}$ will also decrease. Because the power contours in Fig. 1 are convex these changes in π_2 and π_1/π_3 will have opposite effects on the power. However $1-\pi_2$ is relatively a much more slowly varying quantity than $1-2\sqrt{S}/(1+S)$ so that the proportional change in $(n+1)^2(1-2\sqrt{S}/(1+S))$ is greater than that in $M(1-\pi_2)/(n+1)^2$.

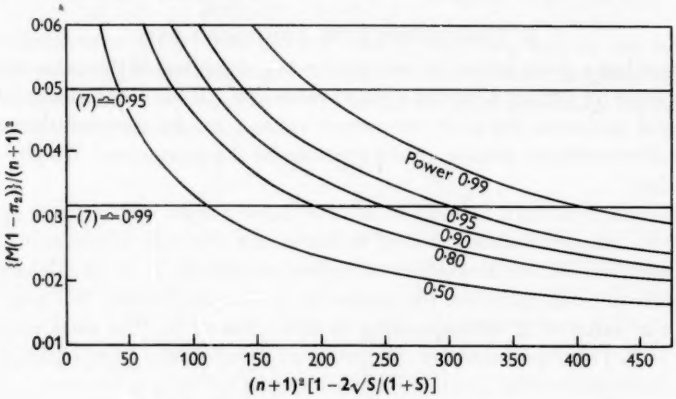


Fig. 1. The approximate power of the test.

We illustrate with two examples:

(i) $P_1 = 0{\cdot}6, \quad P_2' = 0{\cdot}7, \quad P_2 = 0{\cdot}65;$

then $1-\pi_2 = 0{\cdot}46$ changing to $0{\cdot}47$, whereas $1-2\sqrt{S}/(1+S)$ changes from $0{\cdot}0239$ to $0{\cdot}0057$.

(ii) $P_1 = 0{\cdot}95, \quad P_2' = 0{\cdot}98, \quad P_2 = 0{\cdot}97;$

then $1-\pi_2 = 0{\cdot}068$ changing to $0{\cdot}077$, whereas $1-2\sqrt{S}/(1+S)$ changes from $0{\cdot}1026$ to $0{\cdot}0343$.

Suppose we were working with a 5% rejection level of the true null hypothesis, and a power of 0·95. Then for case (1) Fig. 1 shows that, approximately, the power drops to below 0·50 if P_2 is actually 0·65 instead of the required 0·7; and for case (2) the power drops to just above 0·5. One would, of course, naturally expect such a drop since an improvement from the control level 0·6 to 0·65 is not nearly as big as required.

Suppose one were working with the more stringent levels of a 1% rejection level of the true null hypothesis and a power of 0·99. Then for case (1) the power drops below 0·5, and for case (2) to round about 0·7.

Conversely P_2 may be above P_2'.

Suppose we take (iii) $P_1 = 0{\cdot}6$, $P_2' = 0{\cdot}7$, $P_2 = 0{\cdot}75$;

then $1-\pi_2$ changes from 0·46 to 0·45 and $1-2\sqrt{S}/(1+S)$ changes from 0·0239 to 0·1339. For both the weaker and the stronger tests mentioned above, the power apparently increases to far beyond 0·99. Again, take

(iv) $P_1 = 0{\cdot}95$, $P_2' = 0{\cdot}98$, $P_2 = 0{\cdot}99$;

then $1-\pi_2$ changes from 0·068 to 0·059 and $1-2\sqrt{S}/(1+S)$ changes from 0·1026 to 0·2649. For both the weaker and the stronger tests the power apparently increases to well beyond 0·99, although, since S is just over 5 the approximations implicit in the construction of Fig. 1 are probably not very good.

That is to say, in this familiar 'treatment against control' structure, the test appears to behave reasonably well.

I am very grateful to Mrs Eleanor Burrell for doing a substantial part of the calculations.

REFERENCES

ARMITAGE, P. (1957). Restricted sequential procedures. *Biometrika*, **44**, 9–26.

BARTLETT, M. S. (1946). The large sample theory of sequential tests. *Proc. Camb. Phil. Soc.* **42**, 239–44.

BARTLETT, M. S. (1955). *An Introduction to Stochastic Processes*. Cambridge University Press.

EFFECT OF NON-NORMALITY ON THE POWER OF THE ANALYSIS OF VARIANCE TEST

By A. B. L. SRIVASTAVA
Indian Institute of Technology, Kharagpur

1. Introduction

A number of studies have been made of the effect of non-normality on the test functions used for the analysis of variance. The effect on Type I error, i.e. on the distribution under the null hypothesis, was studied by Pearson (1931), Geary (1947) and Gayen (1950), and that on Type II error, i.e. on the power function, has been considered in a study by David & Johnson (1951). Pearson showed that while both 'between-groups' and 'within-groups' mean squares, in the case of non-normal variation, provide unbiased estimates of population variance, they are no longer independently distributed; in fact, their variances and covariance contain a term in λ_4.* However, he reached the conclusion that non-normality would not have a serious effect on the distribution of their ratio w, in large samples. Considering the effect of kurtosis only, Geary (1947) gave an approximate formula for the probability correction for w, based on the large sample assumption. Gayen (1950) derived the distribution of w for non-normal populations specified by the first four terms of the Edgeworth series. He found it to consist of corrective terms in λ_4 and λ_3^2 in addition to the normal theory probability density function of w. Assuming that the observation x_{ij} contains an error term e_{ij} which has a distribution of any form whatever, with all the cumulants existing and varying from group to group, David & Johnson (1951) obtained the moments of the distribution of a certain function of the observations which makes the study of power possible in the non-normal case. They did not, however, make direct use of the distribution of a non-central variance-ratio as is done in the case of the 'normal theory' power function.

Tang (1938) and Patnaik (1949) studied the power of the normal theory χ^2- and F-tests, by deriving the corresponding non-central distributions. The non-central distribution arises when the hypothesis of equal means is not true. The effect of non-normality on the power of the analysis of variance test can likewise be studied by investigating the non-central distribution of the variance-ratio. This has been derived in the present paper, on the assumption that the distribution of the error term is represented by the first four terms of the Edgeworth series. In addition to the normal theory power (Tang, 1938), the corrective terms in λ_3, λ_4 and λ_3^2 have been determined. In deriving conclusions we have kept in view the fact that only the values of λ_3 and λ_4 within certain limits (Barton & Dennis, 1952) can be permitted if the Edgeworth series is to represent a positive definite and unimodal frequency function.

We have considered here the simple case of k groups of n observations. Gayen (1950) obtained the distribution of the variance-ratio assuming unequal number of observations in different groups, but we have avoided this assumption as it would necessitate introduction of more parameters of non-centrality and would make the results more complicated. The results obtained here enable one to calculate the effect of non-normality on the probability of the error of second kind, and hence on the power, when the standard F-table significance levels are used. The difficulty of having a systematic tabulation of the derived expression for

* Here and below we use λ_3 and λ_4 to denote the standardized third and fourth cumulants of the variables.

the power arises from the fact that it contains confluent hypergeometric functions, which do not appear to have been tabulated so far in detail.

2. DISTRIBUTION OF THE NON-CENTRAL VARIANCE-RATIO

Let us consider k groups of observations x_{ij} $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, k)$, and suppose that

$$x_{ij} = A + B_j + e_{ij}, \tag{1}$$

where A is the grand mean, B_j the deviation of the jth group from the grand mean so that $\Sigma B_j = 0$ and e_{ij} represents the random residual distributed non-normally with mean zero, variance σ^2 and standardized third and fourth cumulants $\lambda_3 (= \sqrt{\beta_1})$ and $\lambda_4 (= \beta_2 - 3)$. Assuming the higher order cumulants to be zero, we can express the distribution of the standardized variable $z_{ij} (= e_{ij}/\sigma)$, by the first four terms of the Edgeworth series as

$$f(z) = \phi(z) - \frac{\lambda_3}{6}\phi^{(3)}(z) + \frac{\lambda_4}{24}\phi^{(4)}(z) + \frac{\lambda_3^2}{72}\phi^{(6)}(z), \tag{2}$$

where $\phi(z)$ is the standard normal function, and $\phi^{(\nu)}(z)$ its νth derivative.

Now 'between-groups' and 'within-groups' sums of squares can be written as

$$X = \frac{n}{\sigma^2}\sum_{j=1}^{k}(\bar{x}_j - \bar{x})^2 = n\sum_{j=1}^{k}(\bar{z}_j - \bar{z} + \delta_j)^2, \quad \text{where} \quad \delta_j = B_j/\sigma \tag{3}$$

and

$$Y = \frac{1}{\sigma^2}\sum_{j=1}^{k}\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2 = \sum_{j=1}^{k}\sum_{i=1}^{n}(z_{ij} - \bar{z}_j)^2. \tag{4}$$

In the analysis of variance, as is known, we test the hypothesis H_0 that $\delta_j = 0$ for all j, and use the distribution of

$$w = [(N-k)X]/[(k-1)Y]$$

(with $\delta_j = 0$ for all j) in order to determine the point w_0 such that the probability $P(w \geqslant w_0)$ has a predetermined value α. We reject the hypothesis H_0 only if $w > w_0$, and accept it otherwise. If H_0 is not true, we can determine from the distribution of w (with δ_j's not all zero), the probability of accepting H_0, that is, the probability of the error of second kind.

Now, in order to obtain the power $\beta = 1 - P_{II}$, we first derive the non-central distribution of w. Let us put

$$S_{1j} = \sum_{i=1}^{n} z_{ij} + n\delta_j, \quad S_{2j} = \sum_{i=1}^{n}(z_{ij} - \bar{z}_j)^2 \tag{5}$$

and

$$X_1 = \sum_{j=1}^{k} S_{1j}, \quad X_2 = \frac{1}{n}\sum_{j=1}^{k} S_{1j}^2 \tag{6}$$

so that $X = X_2 - X_1^2/N$, where $N = nk$.

Using the joint distribution of S_{1j} and S_{2j} which is obtained by replacing S_{1j} by $S_{1j} - n\delta_j$ in the result (2·11) of Gayen (1949), the joint characteristic function of X_1, X_2 and Y has been derived as

$$\begin{aligned}
\phi(t_1, t_2, u) = {} & e^{\lambda\tau_{22}}\phi_0(t_1, t_2, u) + \frac{\exp\{-\frac{1}{2}(Nt_1^2 - 2it_2\lambda)/(1-2it_2)\}\,\tau_{22}^2}{(1-2it_2)^{\frac{1}{2}k}\,(1-2iu)^{\frac{1}{2}(N-k)}}\left[\frac{2\lambda_3}{3}\{3\lambda\tau_{12} + 2\mu\tau_{22}\}\right. \\
& + \frac{\lambda_4}{3}\left\{3\lambda\left(\tau_{12}^2 - 1 + \frac{k}{N(1-2it_2)} + \frac{N-k}{N(1-2iu)}\right) + 4\mu\tau_{12}\tau_{22} + 2\nu\tau_{22}^2\right\} \\
& + \frac{\lambda_3^2}{9}\left\{3\lambda\left([N\tau_{12}^4 - 3(N+3)\tau_{12}^2 + 6] + \frac{3}{(1-2it_2)}[(k+3)\tau_{12}^2 - (N+3k)/N]\right.\right. \\
& + \frac{6k}{N(1-2it_2)^2} + \frac{3(N-k)}{N(1-2iu)}(N\tau_{12}^2 - 3) + \frac{3(N-k)}{N(1-2iu)(1-2it_2)} + \frac{3(N-k)}{N(1-2iu)^2}\Big) \\
& + 2\mu\tau_{12}\tau_{22}\left([N\tau_{12}^2 - 3(N+6)] + \frac{3(k+6)}{(1-2it_2)} + \frac{3(N-k)}{(1-2iu)}\right) \\
& \left.\left. + 18\tau_{22}^2(\lambda^2\tau_{12}^2 + 2\nu\tau_{22}) + 24\lambda\mu\tau_{12}\tau_{22}^3 + 8\mu^2\tau_{22}^4\right\}\right].
\end{aligned} \tag{7}$$

where
$$\tau_{12} = \frac{it_1}{(1-2it_2)}, \quad \tau_{22} = \frac{it_2}{(1-2it_2)}; \tag{8}$$

$$\lambda = n\sum_1^k \delta_j^2, \quad \mu = n\sum_1^k \delta_j^3 \quad \text{and} \quad \nu = n\sum_1^k \delta_j^4 * \tag{9}$$

and where $\phi_0(t_1, t_2, u)$ denotes Gayen's (1950) result (2·6)†, in which we have taken $k'^2 = k^2$.

Obtaining the joint frequency function of X_1, X_2 and Y from (7) by Fourier's inversion theorem, and integrating out X_1, we find the joint frequency function of X and Y can be written in the form

$$g(.,X,Y) = g_0(.,X,Y) + \lambda_3 g_{\lambda_3}(.,X,Y) + \lambda_4 g_{\lambda_4}(.,X,Y) + \lambda_3^2 g_{\lambda_3^2}(.,X,Y), \tag{10}$$

the typical term in it being
$$\phi(Y, r_1)\,\psi(X, r_2), \tag{11}$$

where
$$\phi(Y, r_1) = \frac{Y^{\frac{1}{2}r_1 - 1} e^{-\frac{1}{2}Y}}{2^{\frac{1}{2}r_1}\Gamma(\frac{1}{2}r_1)} \tag{12}$$

and
$$\psi(X, r_2) = \frac{e^{-\frac{1}{2}\lambda} e^{-\frac{1}{2}X}}{2^{\frac{1}{2}r_2}} \sum_{j=0}^{\infty} \frac{X^{\frac{1}{2}r_2 + j - 1}\lambda^j}{2^{2j}\Gamma(\frac{1}{2}r_2 + j)\,j!}. \tag{13}$$

We can now deduce the frequency function of $w = [(N-k)X]/[(k-1)Y]$ from (10) in the form
$$g(w) = g_0(w) + \lambda_3 g_{\lambda_3}(w) + \lambda_4 g_{\lambda_4}(w) + \lambda_3^2 g_{\lambda_3^2}(w). \tag{14}$$

Using $p(w; r_1, r_2)$ to denote

$$\sum_{j=0}^{\infty} \frac{e^{-\frac{1}{2}\lambda}(\frac{1}{2}\lambda)^j}{j!\,B(\frac{1}{2}r_1 + j, \frac{1}{2}r_2)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{1}{2}r_1 + j} \frac{w^{\frac{1}{2}r_1 + j - 1}}{\left(1 + \frac{\nu_1}{\nu_2} w\right)^{\frac{1}{2}(r_1 + r_2) + j}}, \tag{15}$$

where $\nu_1 = k-1, \nu_2 = N-k$ are the degrees of freedom of X and Y, we find

$$g_0(w) = p(w; \nu_1, \nu_2), \tag{16}$$

$$g_{\lambda_3}(w) = \tfrac{1}{6}\mu p_3(w; \nu_1, \nu_2) \tag{17}$$

$$\begin{aligned} g_{\lambda_4}(w) = \frac{1}{24N}[&3\{(\nu_1+\nu_2)^2 p(w; \nu_1, \nu_2) - 2\nu_1(\nu_1+\nu_2)\,p(w; \nu_1+2, \nu_2) \\ &+ \nu_1^2 p(w; \nu_1+4, \nu_2) - 2\nu_2(\nu_1+\nu_2)\,p(w; \nu_1, \nu_2+2) + 2\nu_1\nu_2 p(w; \nu_1+2, \nu_2+2) \\ &+ \nu_2^2 p(w; \nu_1, \nu_2+4)\} + 6\lambda\{-(\nu_1+\nu_2)\,p_2(w; \nu_1, \nu_2) + \nu_1 p_2(w; \nu_1+2, \nu_2) \\ &+ \nu_2 p_2(w; \nu_1, \nu_2+2)\} + N\nu p_4(w; \nu_1, \nu_2)], \end{aligned} \tag{18}$$

$$\begin{aligned} g_{\lambda_3^2}(w) = \frac{1}{72N}[&6\{-(\nu_1+\nu_2)(\nu_1+\nu_2-1)\,p(w; \nu_1, \nu_2) + 3\nu_1(\nu_1+\nu_2-1)\,p(w; \nu_1+2, \nu_2) \\ &- 3\nu_1(\nu_1-1)\,p(w; \nu_1+4, \nu_2) + \nu_1(\nu_1-1)\,p(w; \nu_1+6, \nu_2) \\ &+ 3\nu_2(\nu_1+\nu_2-1)\,p(w; \nu_1, \nu_2+2) - 6\nu_1\nu_2 p(w; \nu_1+2, \nu_2+2) \\ &- 3\nu_2(\nu_2-1)\,p(w; \nu_1, \nu_2+4) + 3\nu_1\nu_2 p(w; \nu_1+2, \nu_2+4) \\ &+ \nu_2(\nu_2-\nu_1-1)\,p(w; \nu_1, \nu_2+6)\} + 18\lambda\{(\nu_1+\nu_2-1)\,p_2(w; \nu_1, \nu_2) \\ &- 2(\nu_1-1)\,p_2(w; \nu_1+2, \nu_2) + (\nu_1-1)\,p_2(w; \nu_1+4, \nu_2) - 2\nu_2 p_2(w; \nu_1, \nu_2+2) \\ &+ \nu_2 p_2(w; \nu_1, \nu_2+4)\} + 9(\lambda^2 - N\nu)\{p_4(w; \nu_1, \nu_2) - p_4(w; \nu_1+2, \nu_2)\} \\ &+ N\mu^2 p_6(w; \nu_1, \nu_2)], \end{aligned} \tag{19}$$

* The conventional λ, or non-central parameter, used here should not be confused with the standardized cumulants, λ_3 and λ_4, nor should ν be confused with the degrees of freedom, ν_1 and ν_2, defined below.

† In the first line of formula (2.6) of Gayen (1950), the factor $(1-2it_2)^{\frac{1}{2}}$ in the denominator should read as $(1-2it_2)^{\frac{1}{2}k}$.

where
$$p_r(w; r_1, r_2) = p(w; r_1+2r, r_2) - rp(w; r_1+2r-2, r_2) + \frac{r(r-1)}{2!}p(w; r_1+2r-4, r_2) - \ldots + (-1)^r p(w; r_1, r_2). \quad (20)$$

3. POWER FUNCTION

We are now in a position to evaluate the probability of the error of second kind P_{II} by integrating $g(w)$ over the region of acceptance. If w_0 be the point which determines the critical region, we shall have

$$P_{II}(w_0) = \int_0^{w_0} g(w)\,dw = P_0(w_0) + \lambda_3 P_{\lambda_3}(w_0) + \lambda_4 P_{\lambda_4}(w_0) + \lambda_3^2 P_{\lambda_3^2}(w_0). \quad (21)$$

In this, $P_0(w_0)$ as obtained from (16), is given by

$$P_0(w_0) = \sum_{j=0}^{\infty} \frac{e^{-\frac{1}{2}\lambda}(\frac{1}{2}\lambda)^j}{j!} I_{u_0}(\tfrac{1}{2}\nu_1 + j, \tfrac{1}{2}\nu_2), \quad (22)$$

where $u_0 = (k-1)\,w_0/[(N-k)+(k-1)\,w_0]$ and I_{u_0} denotes the incomplete β-function. This is the normal-theory expression for $P_{II}(w_0)$ and the methods of its evaluation have been discussed by Tang (1938) and Patnaik (1949). Here the expressions for $P_{\lambda_3}(w_0)$, $P_{\lambda_4}(w_0)$ and $P_{\lambda_3^2}(w_0)$ as obtained by integrating (17), (18) and (19), respectively, are expressible in terms of the confluent hypergeometric functions

$$F(a, b, x) = \sum_{j=0}^{\infty} \frac{\Gamma(a+j)\,\Gamma(b)\,x^j}{\Gamma(b+j)\,\Gamma(a)\,j!}.$$

Writing
$$F_{(0)}(a, b, \tfrac{1}{2}\lambda u_0) = \frac{\Gamma(a)}{\Gamma(a-b+1)\,\Gamma(b)} F(a, b, \tfrac{1}{2}\lambda u_0), \quad (23)$$

and
$$\left.\begin{aligned} F_{(r)}(a, b, \tfrac{1}{2}\lambda u_0) &= F_{(0)}(a, b, \tfrac{1}{2}\lambda u_0) - {}^rC_1 u_0 F_{(0)}(a+1, b+1, \tfrac{1}{2}\lambda u_0) \\ &\quad + \ldots + (-1)^r u_0^r F_{(0)}(a+r, b+r, \tfrac{1}{2}\lambda u_0), \\ F_{(r)'}(a, b, \tfrac{1}{2}\lambda u_0) &= F_{(0)}(a, b, \tfrac{1}{2}\lambda u_0) - {}^rC_1 (1-u_0)\, F_{(0)}(a+1, b, \tfrac{1}{2}\lambda u_0) \\ &\quad + \ldots + (-1)^r (1-u_0)^r F_{(0)}(a+r, b, \tfrac{1}{2}\lambda u_0), \end{aligned}\right\} \quad (24)$$

we get
$$P_{\lambda_3}(w_0) = -\tfrac{1}{6}\mu e^{-\frac{1}{2}\lambda} u_0^{\frac{1}{2}\nu_1} (1-u_0)^{\frac{1}{2}\nu_2} F_{(2)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1+2}{2}, \tfrac{1}{2}\lambda u_0\right), \quad (25)$$

$$\begin{aligned} P_{\lambda_4}(w_0) = \frac{1}{8N} e^{-\frac{1}{2}\lambda} u_0^{\frac{1}{2}\nu_1} (1-u_0)^{\frac{1}{2}\nu_2} \Bigg[& \nu_1^2 F_{(1)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1+2}{2}, \tfrac{1}{2}\lambda u_0\right) \\ & - \nu_2^2 F_{(1)'}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1}{2}, \tfrac{1}{2}\lambda u_0\right) - 2\nu_1\nu_2 F_{(1)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1}{2}, \tfrac{1}{2}\lambda u_0\right) \\ & - 2\lambda\left\{\nu_1 F_{(2)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1+2}{2}, \tfrac{1}{2}\lambda u_0\right) - \nu_2 F_{(2)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1}{2}, \tfrac{1}{2}\lambda u_0\right)\right\} \\ & + \frac{N}{3}\nu F_{(3)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1+2}{2}, \tfrac{1}{2}\lambda u_0\right)\Bigg], \end{aligned} \quad (26)$$

$$\begin{aligned} P_{\lambda_3^2}(w_0) = \frac{1}{12N} e^{-\frac{1}{2}\lambda} u_0^{\frac{1}{2}\nu_1} (1-u_0)^{\frac{1}{2}\nu_2} \Bigg[& -\nu_1(\nu_1-1)\, F_{(2)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1+2}{2}, \tfrac{1}{2}\lambda u_0\right) \\ & + \nu_2(\nu_2-1)\, F_{(2)'}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1}{2}, \tfrac{1}{2}\lambda u_0\right) - \nu_1\nu_2\left\{3F_{(2)'}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1+2}{2}, \tfrac{1}{2}\lambda u_0\right)\right. \\ & \left. + F_{(2)'}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1}{2}, \tfrac{1}{2}\lambda u_0\right)\right\} + 3\lambda\left\{(\nu_1-1)\, F_{(3)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1+2}{2}, \tfrac{1}{2}\lambda u_0\right)\right. \\ & \left. + \nu_2 F_{(3)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1-2}{2}, \tfrac{1}{2}\lambda u_0\right)\right\} + \tfrac{3}{2}(\lambda^2 - N\nu)\, F_{(4)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1+2}{2}, \tfrac{1}{2}\lambda u_0\right) \\ & + \frac{N}{6}\mu^2 F_{(5)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1+2}{2}, \tfrac{1}{2}\lambda u_0\right)\Bigg]. \end{aligned} \quad (27)$$

If ν_2 is an even integer, the above expressions can be reduced to the much simpler form of finite series by using the formula

$$F(a+r,a,x) = e^x F(-r,a,-x) = e^x \sum_{i=0}^{r} \frac{B(a,r+1)}{B(a+i,r+1-i)} \frac{x^i}{i!}. \tag{28}$$

For the normal-theory term, $P_0(w_0)$, such an expression in the form of a finite series when ν_2 is even, has already been given by Tang (1938). It is found that by using (28) the $F_{(r)}$ and $F_{(r)'}$ functions occurring in the corrective terms yield expressions which are not difficult to evaluate for ν_2 as large as 20. For example, in order to evaluate $P_{\lambda_3}(w_0)$ we have to calculate the value of $F_{(2)}\{\frac{1}{2}(\nu_1+\nu_2), \frac{1}{2}(\nu_1+2), \frac{1}{2}\lambda u_0)\}$ which can be easily done by using the formula

$$F_{(2)}\left(\frac{\nu_1+\nu_2}{2}, \frac{\nu_1+2}{2}, \tfrac{1}{2}\lambda u_0\right) = e^{\frac{1}{2}\lambda u_0} \sum_{i=0}^{\frac{1}{2}\nu_2-1} \left[\binom{\frac{1}{2}(\nu_1+\nu_2+2)}{\frac{1}{2}(\nu_2-2)-i} u_0^2 - 2\binom{\frac{1}{2}(\nu_1+\nu_2)}{\frac{1}{2}(\nu_2-2)-i} u_0 + \binom{\frac{1}{2}(\nu_1+\nu_2-2)}{\frac{1}{2}(\nu_2-2)-i} \right] \frac{(\frac{1}{2}\lambda u_0)^i}{i!},$$

where

$$\binom{a}{b} = \frac{\Gamma(a+1)}{\Gamma(b+1)\,\Gamma(a-b+1)}.$$

Similar expressions can be worked out for $F_{(1)}$, $F_{(1)'}$, $F_{(2)'}$, $F_{(3)}$, etc., which occur in the equations (26) and (27), in order to simplify the computation of $P_{\lambda_4}(w_0)$ and $P_{\lambda_3^2}(w_0)$ when ν_2 is even.

4. Numerical evaluation of results and discussion

We shall now consider numerical examples to illustrate the nature of the effects of non-normality on the power. It should be mentioned here that the formulae given above are useful only in determining the effects of such non-normality as is not of very serious type. However, a considerable range of values of λ_3 and λ_4 can be covered within the limits given by Barton & Dennis (1952) (see § 1 above).

The values of $P_{\lambda_3}(w_0)$, $P_{\lambda_4}(w_0)$ and $P_{\lambda_3^2}(w_0)$ have been calculated for $\nu_1 = 4, \nu_2 = 10$ and $\nu_1 = 4, \nu_2 = 20$ when the critical region is determined by the 'normal theory' upper 5% values of F and the alternatives are given by $\phi = \sqrt{(\lambda/k)} = 1{\cdot}0,\ 1{\cdot}5,\ 2{\cdot}0,\ 2{\cdot}5$, etc. This critical region is, of course, based on an erroneous F-value; the actual probability of Type I error, however, can be obtained by adding to 0·05 the corrections due to non-normality given by Gayen (1950)*. In the examples considered, since ν_2 is even, it has been possible to use the finite series expansions for calculation. The values are given in Table 1.

It is first necessary to examine what values μ and ν can take when $\lambda = k\phi^2$ has some fixed positive value. We have seen that the null-hypothesis H_0 assumes that δ_j's, the deviations of the group means from the grand mean, are all zero. If differences between groups exist, the δ_j's will not be all zero and their set of values will define an alternative to H_0. In practice, the true values of group differences are not known, and one would depend on the parameters λ, μ and ν determined by δ_j's to define an alternative hypothesis. As indicated by Tang (1938) in an example, two extreme forms of the set of δ_j's under the restriction $\sum_1^k \delta_j = 0$ will be as follows:

(*a*) the k groups form two distinct equal sets (with a single group midway between if k is odd), the effects in one set being equal and opposite in sign to those in the other;

(*b*) of the k groups, $k-1$ are quite similar and only one is divergent, with its effect counterbalancing the rest.

* There are a few slips in Gayen's (1950, pp. 242) formulae (2·30) to (2·33 *bis*). There should be a *minus* sign before λ_4 in (2·30) instead of the *plus* sign and in all the formulae from (2·31) to (2·33 *bis*) x_0 should be replaced by $1-x_0$.

For a fixed λ, in the case (a) the value of μ will be zero and that of ν will be minimum $(=\lambda^2/N$ or $k\lambda^2/[N(k-1)]$ according as k is even or odd). In the case (b), μ will have an extreme positive or negative value $(=\pm(k-2)\lambda^{\frac{3}{2}}/[N(k-1)]^{\frac{1}{2}})$ and ν will be maximum $(=(k^2-3k+3)\lambda^2/[N(k-1)])$. In practice the situation may be anywhere between these two extreme cases.

Table 1. *Showing the values of the corrective functions for* $\alpha = 0{\cdot}05$ *required in equation* (21) *when* (*i*) $\nu_1 = 4$, $\nu_2 = 10$ *and* (*ii*) $\nu_1 = 4$, $\nu_2 = 20$

D.F.	ϕ	$P_0(w_0)$	$P_{\lambda_3}(w_0)$	$P_{\lambda_4}(w_0)$	$P_{\lambda_3^2}(w_0)$
$\nu_1 = 4$	0·0	0·950	0	0·00376	−0·00429
$\nu_2 = 10$	1·0	0·738	0·0010750 μ	−0·01273 − 0·0000744 ν	0·00480 − 0·0000213 ν + 0·00000604 μ^2
	1·5	0·449	0·0002396 μ	−0·01916 − 0·0000489 ν	0·00300 + 0·0000422 ν − 0·00000007 μ^2
	2·0	0·179	−0·0001670 μ	−0·00156 − 0·0000045 ν	−0·00390 + 0·0000157 ν − 0·00000067 μ^2
	2·5	0·043	−0·0001089 μ	0·00944 + 0·0000047 ν	−0·00174 − 0·0000008 ν − 0·00000008 μ^2
	3·0	0·006	−0·0000239 μ	0·00495 + 0·0000015 ν	0·00085 − 0·0000010 ν + 0·00000002 μ^2
$\nu_1 = 4$	0·0	0·950	0	0·00241	−0·00096
$\nu_2 = 20$	1·0	0·681	0·0018003 μ	−0·01122 − 0·0001695 ν	0·00273 − 0·0000038 ν + 0·00001832 μ^2
	1·5	0·341	0·0000834 μ	−0·00929 − 0·0000818 ν	−0·00159 + 0·0001133 ν − 0·00000190 μ^2
	2·0	0·093	−0·0003425 μ	0·00050 + 0·0000083 ν	−0·00182 + 0·0000169 ν − 0·00000148 μ^2
	2·5	0·012	−0·0000978 μ	0·00458 + 0·0000074 ν	0·00084 − 0·0000083 ν + 0·00000006 μ^2

N.B. The parameter λ, μ and ν have been defined in equations (9).

In Table 2, the values of $P_{\lambda_3}(w_0)$, $P_{\lambda_4}(w_0)$ and $P_{\lambda_3^2}(w_0)$ (of Table 1) are given for the following cases: (a) $\mu = 0$, ν maximum; (b_1) μ maximum positive, ν minimum; and (b_2) μ maximum negative, ν minimum. The sign of μ, which distinguishes case (b_2) from (b_1), only affects $P_{\lambda_3}(w_0)$. Where there are two signs with a value for case (b), the upper sign relates to (b_1). The values tabled are expected to be correct to five places of decimals.

In the examples, the effects of λ_3, λ_4 and λ_3^2 are observed generally in the third place of decimals but sometimes in the second place. The effect of λ_4 is, in general, higher than that of λ_3 or λ_3^2. Also it appears that the values of $P_{\lambda_4}(w_0)$, as well as of $P_{\lambda_3^2}(w_0)$, are not much different in the cases (a) and (b).

We must now compute the power of the test in the case of non-normal populations with given values of λ_3 and λ_4. In Table 3, such comparison is given for the examples considered above. Here the two entries before any value of ϕ in a column with $\lambda_3 = 0$ correspond to the cases (a) and (b), respectively, and the three entries for $\lambda_3 \neq 0$, correspond to the casess (a), (b_1) and (b_2). It is to be noted that the values for the cases (b_1) and (b_2) will interchange as λ_3 changes sign.

Table 2. *Giving the values of* $P_{\lambda_3}(w_0)$, $P_{\lambda_4}(w_0)$ *and* $P_{\lambda_3^2}(w_0)$ *(of Table* 1*) in the cases* (a), (b_1) *and* (b_2)

ϕ	Case	$\nu_1 = 4, \nu_2 = 10$			$\nu_1 = 4, \nu_2 = 20$		
		$P_{\lambda_3}(w_0)$	$P_{\lambda_4}(w_0)$	$P_{\lambda_3^2}(w_0)$	$P_{\lambda_3}(w_0)$	$P_{\lambda_4}(w_0)$	$P_{\lambda_3^2}(w_0)$
0·0		0	0·00376	−0·00429	0	0·00241	−0·00096
1·0	(a)	0	−0·01289	0·00476	0	−0·01143	0·00273
	(b)	±0·00465	−0·01313	0·00479	±0·00604	−0·01177	0·00293
1·5	(a)	0	−0·01968	−0·00345	0	−0·00981	−0·00087
	(b)	±0·00350	−0·02050	−0·00415	±0·00094	−0·01064	0·00003
2·0	(a)	0	−0·00171	−0·00338	0	−0·00067	−0·00148
	(b)	∓0·00579	−0·00195	−0·00334	∓0·00919	0·00093	−0·00201
2·5	(a)	0	0·00982	−0·00181	0	0·00494	0·00058
	(b)	∓0·00737	0·01043	−0·00228	∓0·00502	0·00552	0·00032
3·0	(a)	0	0·00520	0·00068	—	—	—
	(b)	∓0·00279	0·00561	0·00068	—	—	—

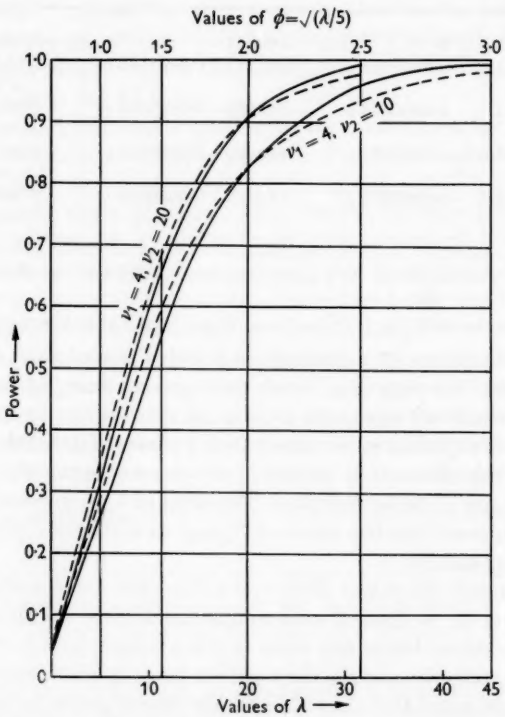


Fig. 1. Showing comparison of powers. —, normal; - - - -, non-normal ($\lambda_3 = 0$, $\lambda_4 = 2{\cdot}4$).

It is again seen that when $\lambda_3 = 0$ there is practically no difference between the values of power in the situations (a) and (b). When $\lambda_3 \neq 0$, in the three situations (a), (b_1) and (b_2) that arise, the powers differ to some extent though not very significantly from the practical point of view. The situation (a), in this case, gives rise to values which are generally intermediate between those of (b_1) and (b_2).

Table 3. *Showing comparison of powers of the F-test in normal and non-normal cases when using the normal theory with 5 % significance level*

ϕ	Symmetrical distributions, $\lambda_3 = 0{\cdot}0$						Skew distributions, $\lambda_3 = 0{\cdot}5$				$\lambda_3 = 0{\cdot}7$
	$\lambda_4 = -1{\cdot}0$	$\lambda_4 = 0{\cdot}0$	$\lambda_4 = 0{\cdot}5$	$\lambda_4 = 1{\cdot}0$	$\lambda_4 = 2{\cdot}0$	$\lambda_4 = 2{\cdot}4$	$\lambda_4 = -1{\cdot}0$	$\lambda_4 = 0{\cdot}0$	$\lambda_4 = 0{\cdot}5$	$\lambda_4 = 2{\cdot}0$	$\lambda_4 = 2{\cdot}4$
Case (i) $\nu_1 = 4$, $\nu_2 = 10$											
0·0	0·054	**0·050**	0·048	0·046	0·042	0·041	0·055	0·051	0·049	0·044	0·043
1·0	0·249	**0·262**	0·268	0·275	0·288	0·290	0·248	0·261	0·267	0·287	0·291
	·249	—	·269	·275	·288	·294	·245	·258	·265	·285	·288
							·250	·263	·270	·289	·294
1·5	0·531	**0·551**	0·561	0·571	0·590	0·598	0·532	0·552	0·562	0·591	0·600
	·531	—	·561	·572	·592	·600	·530	·550	·566	·591	·600
							·533	·554	·564	·595	·605
2·0	0·819	**0·821**	0·822	0·823	0·824	0 825	0·820	0·822	0·823	0·825	0·827
	·819	—	·822	·823	·825	·826	·823	·825	·826	·829	·831
							·817	·819	·820	·823	·823
2·5	0·967	**0·957**	0·952	0·947	0·937	0·933	0·967	0·957	0·953	0·938	0·934
	·967	—	·952	·947	·936	·932	·972	·961	·956	·940	·938
							·964	·954	·949	·933	·928
3·0	0·999	**0·994**	0·991	0·989	0·984	0·982	0·999	0·994	0·991	0·983	0·981
	—	—	·991	·988	·983	·981	·999	·995	·992	·984	·982
							·998	·992	·990	·981	·978
Case (ii) $\nu_1 = 4$, $\nu_2 = 20$											
0·0	0·052	**0·050**	0·049	0·048	0·045	0·044	0·053	0·050	0·049	0·045	0·045
1·0	0·308	**0·319**	0·325	0·330	0·342	0·346	0·307	0·318	0·324	0·341	0·345
	·307	—	·325	·331	·343	·347	·303	·315	·321	·339	·342
							·310	·321	·327	·345	·350
1·5	0·649	**0·659**	0·664	0·669	0·679	0·683	0·649	0·659	0·664	0·679	0·683
	·648	—	·664	·670	·680	·685	·648	·658	·664	·680	·684
							·649	·659	·665	·681	·685
2·0	0·908	**0·907**	0·907	0·906	0·906	0·905	0·908	0·907	0·907	0·906	0·906
	·908	—	·907	·906	·905	·905	·913	·912	·912	·910	·912
							·904	·903	·902	·901	·899
2·5	0·993	**0·988**	0·986	0·983	0·978	0·976	0·993	0·988	0·985	0·978	0·976
	·994	—	·985	·982	·977	·975	·996	·990	·988	·979	·978
							·991	·985	·983	·974	·971

Power curves relating to the symmetric lepto-kurtic population with $\lambda_4 = 2{\cdot}4$ ($\beta_2 = 5{\cdot}4$) have been drawn in Fig. 1 along with the 'normal theory' power curves for the cases $\nu_1 = 4, \nu_2 = 10$ and $\nu_1 = 4, \nu_2 = 20$. This diagram illustrates the nature and magnitude of the effect of such a departure from normality in the distribution of the random residuals. Power curves in the case of the other non-normal populations considered in Table 3 would differ from normal theory to a less extent.

5. Conclusion

Obtaining the expressions for the power function of the analysis of variance test for a one-way classification in the case where the sampled population is represented by the first four terms of an Edgeworth series, we have considered certain numerical examples to see what is the nature of effects of non-normality on the power. The populations considered are only moderately non-normal, as the terms higher than the fourth in the Edgeworth series are assumed to be zero. Also the values of λ_3 and λ_4 have not been allowed to exceed the Barton & Dennis (1952) limits, ensuring thereby that the frequency function represented is positive definite and unimodal.

It is found that the effect of skewness is not much on the power of the analysis of variance test, at least when the values of λ_3 are confined to Barton & Dennis limits. In practice, however, higher values of λ_3 may occur and then the effect of skewness may be somewhat larger and comparable to that of kurtosis which is, in general, high. The presence of a fair degree of kurtosis, as is not uncommon in practice, leads to a noticeable change in the power curve particularly in the case of small samples. But a small departure from normality in respect of kurtosis (say, of the order $\lambda_4 = \pm 0{\cdot}5$) again does not cause any significant deviation in the power. When the population is leptokurtic ($\lambda_4 > 0$), the power increases in the beginning (for example, up to the point for which the power is approximately 0·8 in the case when $\nu_1 = 4, \nu_2 = 10$ (Fig. 1)), but in the region of very high power, it subsequently decreases in comparison with the normal-theory power. The reverse happens when λ_4 is negative.

The above conclusions derived from numerical results are expected to be valid in general. There is a good indication that the effect of non-normality on the power diminishes with increasing sample size, as expected. In practice the effect of kurtosis is likely to be more on the power than that of skewness. A general conclusion in this regard cannot, however, be given, for it is difficult to take into account all the ways in which the values of λ_3 and λ_4 of the distributions met in practice will vary. But on the whole it may be said that from the practical point of view, the effect of non-normality on power will not be of much consequence in the case of near-normal populations.

My thanks are due to Dr A. K. Gayen for his help and guidance in the course of preparation of this work. I am also grateful to Prof. E. S. Pearson for certain comments and suggestions leading to some improvement of the paper.

REFERENCES

Barton, D. E. & Dennis, K. E. (1952). *Biometrika*, **39**, 425–7.
David, F. N. & Johnson, N. L. (1951). *Biometrika*, **38**, 43–57.
Gayen, A. K. (1949). *Biometrika*, **36**, 353–69.
Gayen, A. K. (1950). *Biometrika*, **37**, 236–55.
Geary, R. C. (1947). *Biometrika*, **34**, 209–42.
Patnaik, P. B. (1949). *Biometrika*, **36**, 202–32.
Pearson, E. S. (1931). *Biometrika*, **23**, 114–33.
Tang, P. C. (1938). *Statist. Res. Mem.* **2**, 126–57.

THE z-TEST AND SYMMETRICALLY DISTRIBUTED RANDOM VARIABLES

BY F. N. DAVID*

University College London and Statistical Laboratory, University of California, Berkeley

Some years ago in two papers (David & Johnson 1951*b*, 1952) the exact cumulants and cross-cumulants of the distributions of the sums of squares usually associated as test criteria for the general linear hypothesis were given under extremely wide assumptions. It had been the intention, at this stage, to substitute in the given formulae the expressions associated with each of the analysis of variance set-ups and thus, by taking appropriate numerical values, to obtain some idea of the effect of non-normality, with and without unequal group frequencies, on the F-test as commonly applied. The arithmetic, however, proved formidable and an algebraic alternative was sought for. This algebraic alternative has proved of use in other problems not associated with the general linear hypothesis. It is proposed therefore to summarize it here and to illustrate its use on the F-test of the one-way classification, the non-normality of which has previously been discussed by Pearson (1931), David & Johnson (1951*a*), and Gayen (1950). It is hoped to present similar formulae using the cumulants for the general linear hypothesis sums of squares in a further publication.

Assume k groups containing $n_1, n_2, \ldots, n_k$ observations with

$$\sum_{t=1}^{k} n_t = n.$$

In commonly used notation, the one-way classification leads to the identity

$$\sum_{t=1}^{k} n_t(\bar{x}_{t.}-\bar{x}_{..})^2+\sum_{t=1}^{k}\sum_{i=1}^{n_t}(x_{ti}-\bar{x}_{t.})^2=\sum_{t=1}^{k}\sum_{i=1}^{n_t}(x_{ti}-\bar{x}_{..})^2.$$

Usually the x's are assumed each to follow the same normal distribution. We shall assume that each has the same symmetrical distribution. There are no real difficulties introduced by still further generalizing the situation and assuming for each group that x has a different asymmetrical distribution. The purpose here however is one of method and the symmetrical distribution leads to shortening of the expressions for cumulants. If

$$S_1=\sum_{t=1}^{k} n_t(\bar{x}_{t.}-\bar{x}_{..})^2 \quad \text{and} \quad S_2=\sum_{t=1}^{k}\sum_{i=1}^{n_t}(x_{ti}-\bar{x}_{t.})^2,$$

then we may write down the cumulants and cross-cumulants of S_1 and S_2 from those already given by David & Johnson (1951*a*). We have that

$$\mathscr{E}(S_1)=(k-1)\kappa_2=f_1\kappa_2, \quad \mathscr{E}(S_2)=(n-k)\kappa_2=f_2\kappa_2.$$

We define

$$\gamma_r=\frac{\kappa_{r+2}}{\kappa_2^{\frac{1}{2}r+1}},$$

$$K(S_1^d S_2^e)=\frac{\kappa(S_1^d S_2^e)}{(\kappa(S_1))^d(\kappa(S_2))^e}$$

and

$$V_{abc}=\frac{n^b}{(n-k)^a k^{b+1-a-c}(k-1)^c}\sum_{t=1}^{k}\frac{(n_t-1)^a}{n_t^b}\left(1-\frac{n_t}{n}\right)^c.$$

* With the partial support of the Office of Naval Research, U.S. Navy.

It will be noticed that $V_{abc} = 1$ when $n_t = n/k, t = 1, 2, \ldots, k$. From the printed cumulants we get, putting the odd gammas zero, and rearranging

$$K(S_1^2) = \gamma_2 \frac{V_{012}}{n} + \frac{2}{f_1}, \quad K(S_1 S_2) = \gamma_2 \frac{V_{111}}{n}, \quad K(S_2^2) = \gamma_2 \frac{V_{210}}{n} + \frac{2}{f_2},$$

$$K(S_1^3) = \gamma_4 \frac{V_{023}}{n^2} + 12\gamma_2 \frac{V_{012}}{nf_1} + \frac{8}{f_1^2},$$

$$K(S_1^2 S_2) = \gamma_4 \frac{V_{122}}{n^2} + 4\gamma_2 \frac{V_{111}}{nf_1},$$

$$K(S_1 S_2^2) = \gamma_4 \frac{V_{221}}{n^2} + 4\gamma_2 \frac{V_{111}}{nf_2},$$

$$K(S_2^3) = \gamma_4 \frac{V_{320}}{n^2} + 12\gamma_2 \frac{V_{210}}{nf_2} + \frac{8}{f_2^2}.$$

$$K(S_1^4) = \gamma_6 \frac{V_{034}}{n^3} + 24\gamma_4 \frac{V_{023}}{n^2 f_1} + \gamma_2^2 \frac{32V_{023}}{n^2 f_1} + 144\gamma_2 \frac{V_{012}}{nf_1^2} + \frac{48}{f_1^3},$$

$$K(S_1^3 S_2) = \gamma_6 \frac{V_{133}}{n^3} + 12\gamma_4 \frac{V_{122}}{n^2 f_1} + \gamma_2^2 \frac{12V_{122}}{n^2 f_1} + 24\gamma_2 \frac{V_{111}}{nf_1},$$

$$K(S_1^2 S_2^2) = \gamma_6 \frac{V_{232}}{n^3} + 4\gamma_4 \frac{V_{111}}{nf_1 f_2} + \gamma_2^2 \left(\frac{12V_{112}}{nf_1 f_2} + \frac{4V_{211}}{nf_1 f_2} \right) + \frac{16\gamma_2 V_{111}}{nf_1 f_2},$$

$$K(S_1 S_2^3) = \gamma_6 \frac{V_{331}}{n^3} + 12\gamma_4 \frac{V_{221}}{n^2 f_2} + 4\gamma_2^2 \left(\frac{7V_{221}}{n^2 f_2} - \frac{2V_{111}}{nf_2^2} \right) + 24\gamma_2 \frac{V_{111}}{nf_2^2},$$

$$K(S_2^4) = \gamma_6 \frac{V_{430}}{n^3} + 24\gamma_4 \frac{V_{320}}{n^2 f_2} + 8\gamma_2^2 \left(\frac{6V_{320}}{n^2 f_2} - \frac{3V_{210}}{nf_2^2} + \frac{1}{f_2^3} \right) + 144\gamma_2 \frac{V_{210}}{nf_2^2} + \frac{48}{f_2^3}.$$

These expressions are all exact except that, in the first three cumulants of the fourth order, terms in γ_2^2 which were of order n^{-4} and n^{-5} have been omitted.

Fisher and Cornish in their paper dealing with the inversion of Edgeworth's expansion showed that the cumulative function of

$$z = \tfrac{1}{2} \log F$$

could be adequately represented in the normal case by the expansion which they obtained. Thus if the cumulants of normal z, say z_N are $\kappa_{1N}, \kappa_{2N}, \ldots, \kappa_{rN}, \ldots$ and the normal curve generating the expansion is allowed to have the same mean and variance as z_N, we have that the standardized deviation from the unit normal deviate ξ, say, of the percentage point, is given by

$$\frac{z_N - \kappa_{1N}}{(\kappa_{2N})^{\frac{1}{2}}} - \xi = \tfrac{1}{6} l_3 (\xi^2 - 1) + \tfrac{1}{24} l_4 (\xi^3 - 3\xi) - \tfrac{1}{36} l_3^2 (2\xi^3 - 5\xi) + \ldots,$$

where

$$l_r = \frac{\kappa_{rN}}{\kappa_{2N}^{\frac{1}{2}r}} \quad (r > 2).$$

Provided the moments of z are not greatly affected by non-normality or by unequal group frequencies there is no reason why a Fisher–Cornish expansion should not be adequate in the non-normal case. We shall assume therefore that the departures are of such order for this to be true. If $L_1, L_2, \ldots$ refer to non-normal z, z_A (say), then let

$$L_1 = \frac{\kappa_{1A} - \kappa_{1N}}{(\kappa_{2N})^{\frac{1}{2}}}, \quad L_2 = \frac{\kappa_{2A} - \kappa_{2N}}{\kappa_{2N}}, \quad L_r = \frac{\kappa_{rA}}{(\kappa_{2N})^{\frac{1}{2}r}} \quad (r > 2).$$

In this case the Fisher–Cornish expansion is

$$\frac{z_A - \kappa_{1N}}{\sqrt{\kappa_{2N}}} - \xi = L_1 + \tfrac{1}{6}L_3(\xi^2 - 1) + \tfrac{1}{2}L_2\xi + \tfrac{1}{24}L_4(\xi^3 - 3\xi) - \tfrac{1}{36}L_3^2(2\xi^3 - 5\xi) + \ldots.$$

Hence by subtraction

$$\frac{z_A - z_N}{\sqrt{\kappa_{2N}}} = \frac{\kappa_{1A} - \kappa_{1N}}{\sqrt{\kappa_{2N}}} + \frac{\xi}{2}\frac{(\kappa_{2A} - \kappa_{2N})}{\kappa_{2N}} + \frac{(\xi^2 - 1)}{6}\left(\frac{\kappa_{3A} - \kappa_{3N}}{\kappa_{2N}^{\frac{3}{2}}}\right) + \frac{(\xi^3 - 3\xi)}{24}\left(\frac{\kappa_{4A} - \kappa_{4N}}{\kappa_{2N}^2}\right) - \tfrac{1}{36}(2\xi^3 - 5\xi)\left[\frac{\kappa_{3A}^2 - \kappa_{3N}^2}{\kappa_{2N}^3}\right] + \ldots. \qquad (1)$$

If desired the cumulants of z_N can be found from the differentiation of the cumulant generating function (see, for example, Wishart (1947))

$$\mathscr{K} = \log\left(\frac{f_1 + it - 2}{2}\right)! - \log\left(\frac{f_1 - 2}{2}\right)! - \tfrac{1}{2}it\log(\tfrac{1}{2}f_1) - \log\left(\frac{f_2 + it - 2}{2}\right)! + \log\left(\frac{f_2 - 2}{2}\right)! + \tfrac{1}{2}it\log(\tfrac{1}{2}f_2),$$

for the polygamma functions are well tabulated, but this is not really necessary for the difference. We also note that z_N is the tabulated percentage point of z, given, for instance, in Fisher and Yates' Tables. z_A can therefore be obtained provided the differences of the cumulants of z_A and z_N are known to a sufficient order of approximation. κ_{2N} is of order n^{-1} and l_r and L_r are of order $n^{1-\frac{1}{2}r}$ when the group frequencies are equal.

By definition

$$2z = \log F = \log(S_1/f_1) - \log(S_2/f_2)$$

or

$$2z = \log\left(1 + \frac{S_1 - \mathscr{E}(S_1)}{\mathscr{E}(S_1)}\right) - \log\left(1 + \frac{S_2 - \mathscr{E}(S_2)}{\mathscr{E}(S_2)}\right).$$

Write δ_1 and δ_2 for reduced, standardized, S_1 and S_2, respectively, and we have

$$2z = (\delta_1 - \delta_2) - \tfrac{1}{2}(\delta_1^2 - \delta_2^2) + \tfrac{1}{3}(\delta_1^3 - \delta_2^3) - \tfrac{1}{4}(\delta_1^4 - \delta_2^4) + \ldots$$

δ_1 and δ_2 are the two quantities the cumulants of which were given earlier. Taking expectations and remembering that those terms remaining when the γ's are put equal to zero are those belonging to twice the mean value of normal z we have

$$\kappa_{1A} - \kappa_{1N} = -\frac{\gamma_2}{4n}(V_{012} - V_{210}) + \left\{\frac{\gamma_4}{6n^2}(V_{023} - V_{320}) - \frac{3}{8}\frac{\gamma_2^2}{n^2}(V_{012}^2 - V_{210}^2) + \frac{\gamma_2}{2n}\left(\frac{V_{012}}{f_1} - \frac{V_{210}}{f_2}\right)\right\} + \ldots$$

to order n^{-2}. Following the usual procedure we obtain for the higher cumulants,

$$\kappa_{2A}-\kappa_{2N}=\frac{\gamma_2}{4n}(V_{012}-2V_{111}+V_{210})+\left\{-\frac{\gamma_4}{4n^2}(V_{023}-V_{122}-V_{221}+V_{320})\right.$$
$$+\frac{\gamma_2^2}{4n^2}(\tfrac{5}{2}V_{012}^2+\tfrac{5}{2}V_{210}^2-2V_{111}V_{012}-2V_{111}V_{210}-V_{111}^2)$$
$$\left.-\frac{\gamma_2}{8n}\left(V_{012}\left(\frac{3}{f_1}+\frac{1}{f_2}\right)+V_{210}\left(\frac{1}{f_1}+\frac{3}{f_2}\right)\right)\right\}+\ldots,$$

$$\kappa_{3A}-\kappa_{3N}=\frac{\gamma_4}{8n^2}(V_{023}-3V_{122}+3V_{221}-V_{320})-\frac{3}{8}\frac{\gamma_2^2}{n^2}(V_{012}-V_{210})(V_{012}+V_{210}-2V_{111})+\ldots,$$

$$\kappa_{4A}-\kappa_{4N}=\frac{\gamma_6}{16n^3}(V_{034}-4V_{133}+6V_{232}-4V_{331}+V_{430})$$
$$+\frac{\gamma_4\gamma_2}{8n^3}(V_{023}(-5V_{012}+6V_{111}-V_{210})+V_{320}(-5V_{210}+6V_{111}-V_{120})$$
$$+V_{122}(9V_{012}-3V_{210}-6V_{111})+V_{221}(9V_{210}-3V_{012}-6V_{111}))$$
$$+\frac{\gamma_2^3}{4n^3}(5V_{012}^3+5V_{210}^3+2V_{111}^3+3V_{111}^2(V_{012}+V_{210})$$
$$-12V_{111}(V_{012}^2+V_{210}^2)+6V_{111}V_{210}V_{012})$$
$$+\frac{\gamma_4}{4n^2}\left(V_{122}\left(\frac{3}{f_1}-\frac{6}{f_2}\right)+V_{221}\left(\frac{3}{f_2}-\frac{6}{f_1}\right)+V_{111}\frac{6n}{f_1f_2}\right)$$
$$+\frac{\gamma_2^2}{2n^2}\left(-\frac{3}{f_1}(V_{012}-V_{111})^2-\frac{3}{f_2}(V_{210}-V_{111})^2+\frac{1}{f_1}\left(4V_{023}-6V_{122}+9V_{112}\frac{n}{f_2}\right)\right.$$
$$\left.+\frac{1}{f_2}\left(6V_{320}-14V_{221}+3V_{211}\frac{n}{f_1}\right)+\frac{n}{f_2^2}\left(4V_{111}-3V_{210}+\frac{n}{f_2}\right)\right).$$

The second and third cumulants are of order n^{-2} and the fourth is of order n^{-3}.

If the numbers in the groups are all equal, the V's are all unity and considerable simplification results. We have

$$\kappa_{1A}-\kappa_{1N}=\frac{\gamma_2}{2n}\left(\frac{1}{f_1}-\frac{1}{f_2}\right)+\ldots$$
$$\kappa_{2A}-\kappa_{2N}=-\frac{\gamma_2}{2n}\left(\frac{1}{f_1}+\frac{1}{f_2}\right)+\ldots$$
$$\kappa_{3A}-\kappa_{3N}=0+\ldots$$
$$\kappa_{4A}-\kappa_{4N}=\frac{\gamma_4}{4n^2}\left(-\frac{3}{f_1}-\frac{3}{f_2}+\frac{6n}{f_1f_2}\right)+\frac{\gamma_2^2}{2n^2}\left(-\frac{2}{f_1}-\frac{8}{f_2}+\frac{12n}{f_1f_2}+\frac{n}{f_2^2}+\frac{n^2}{f_2^3}\right)+\ldots.$$

It is possible to rewrite the moments in a slightly different way which will bring out more strongly the variability of the group frequencies and the effect which it has on the moments of z. Let, for $i=1,2,\ldots,k$,

$$n_i=\frac{n}{k}+\Delta_i;$$

then
$$\sum_{i=1}^{k}\Delta_i^2=k\,\mathrm{var}\,(n_i)=kv \quad \text{(say)}.$$

We take the expression for V_{abc}, substitute for n_i, and expand to get

$$V_{abc}-1=v\left[{}^{a}C_2\frac{k^2}{f_2^2}-\frac{ack^2}{nf_1f_2}+{}^{c}C_2\frac{k^2}{n^2f_1^2}+{}^{b+1}C_2\frac{k^2}{n^2}-\frac{abk^2}{nf_2}+\frac{bck^2}{n^2f_1}\right]$$
$$+\mu_3(n)\left[{}^{a}C_3\frac{k^3}{f_2^3}-{}^{a}C_2\,{}^{c}C_1\frac{k^3}{nf_1f_2^2}+{}^{c}C_2\,{}^{a}C_1\frac{k^3}{n^2f_1^2f_2}-{}^{c}C_3\frac{k^3}{n^3f_1^3}-{}^{b+2}C_3\frac{k^3}{n^3}+{}^{b+1}C_2\,{}^{a}C_1\frac{k^3}{n^2f_2}\right.$$
$$\left.-{}^{b+1}C_2\,{}^{c}C_1\frac{k^3}{n^2f_1}-{}^{a}C_2\,{}^{b}C_1\frac{k^3}{nf_2^2}+{}^{a}C_1\,{}^{b}C_1\,{}^{c}C_1\frac{k^3}{n^2f_1f_2}-{}^{b}C_1\,{}^{c}C_2\frac{k^3}{n^3f_1^2}\right]+\ldots.$$

For example
$$V_{111}-1=-\frac{vk^4}{n^2f_1f_2}+\frac{k^5}{n^3f_1f_2}\mu_3(n).$$

If $k=5$ and $n=20$ this gives us

$$V_{111}-1=-\frac{5}{192}v+\frac{5}{768}\mu_3(n)$$

which suggests that for all but widely different group frequencies the first term involving v will be adequate, and this is so in fact. We therefore substitute for V_{abc} and collect terms which enables the effect of the unequal group frequencies, as described by their variance, to be seen immediately.

Let
$$A_r=k^r\left(\frac{1}{f_1^r}-\frac{1}{f_2^r}\right)$$
$$B_r=k^r\left(\frac{1}{f_1^r}+\frac{1}{f_2^r}\right).$$

Then on substitution for V_{abc} we have

$$\kappa_{1A}-\kappa_{1N}=-\frac{\gamma_2k^2A_2v}{4n^3}+\left\{\frac{\gamma_2A_1}{2nk}+\frac{v}{2}\frac{k^2}{n^4}\left(\gamma_4A_2-\tfrac{3}{2}\gamma_2^2A_2+\gamma_2\frac{n}{k}A_3\right)-\tfrac{3}{8}v^2\frac{k^4}{n^6}\gamma_2^2A_4\right\}+\ldots$$

$$\kappa_{2A}-\kappa_{2N}=\frac{\gamma_2k^2B_1^2v}{4n^3}+\left\{-\frac{\gamma_2B_1}{2nk}+\frac{vk^2B_1^2}{4n^4}(-2\gamma_4+3\gamma_2^2)+\frac{vk}{4n^3}\gamma_2(B_1B_2-A_1A_2)\right.$$
$$\left.+\frac{v^2k^4\gamma_2^2}{4n^6}\left(\tfrac{11}{4}B_1^2-\frac{9k^2}{f_1f_2}B_1^2-\frac{A_2^2}{4}\right)\right\}+\ldots$$

$$\kappa_{3A}-\kappa_{3N}=-\tfrac{3}{8}v^2\frac{k^4}{n^6}B_1^2A_2\gamma_2^2+\ldots$$

$$\kappa_{4A}-\kappa_{4N}=-\frac{3}{4}\frac{\gamma_4\gamma_2}{n^3}\frac{k^4}{n^4}B_1^4v^2$$
$$+\frac{\gamma_2^3}{4n^3}\left(\frac{3v^2k^4}{n^4}\left(B_2^2+\frac{8k^2}{f_1f_2}B_2\right)+\frac{v^3k^6}{n^6}\left(5B_6-\frac{8k^6}{f_1^3f_2^3}+\frac{3k^4}{f_1^2f_2^2}B_2+\frac{12k^2}{f_1f_2}B_4\right)\right)$$
$$+\frac{\gamma_4}{4n^2}\left(-3\left(\frac{1}{f_1}+\frac{1}{f_2}\right)+\frac{6n}{f_1f_2}+\frac{kv}{n^2}\left(3B_3-\frac{6nk^3}{f_1^2f_2^2}\right)\right)$$
$$+\frac{\gamma_2^2}{2n^2}\left(-\frac{2}{f_1}-\frac{8}{f_2}+\frac{12n}{f_1f_2}+\frac{n}{f_2^2}+\frac{n^2}{f_2^3}+v\left(\frac{4k}{n^2}B_1^3+\frac{k^4}{n^2}\left(\frac{2}{f_1^3}+\frac{16}{f_1f_2^2}\right)\right.\right.$$
$$\left.\left.+\frac{k^2}{f_1f_2}\left(-\frac{12}{f_1f_2}-\frac{12k}{f_1f_2}+\frac{k^2}{n}\left(\frac{9}{f_1^2}+\frac{3}{f_2^2}\right)\right)+v^2\left(-\frac{3k^3}{n^4}B_1^2B_3\right)\right)\right).$$

When $v=0$, that is when all the numbers in the groups are equal, the cumulants reduce to those previously noted.

We consider the case of equal group frequencies first. Refer to equation (1), substitute in the cumulant differences and collect terms of like order. The difference between the tabulated value of z and that of non-normal z is to order $n^{-\frac{3}{2}}$ and for k reasonably large

$$-\frac{1}{\sqrt{2}}\frac{\gamma_2}{n}\left(\frac{B_1}{k}\right)^{\frac{1}{2}}\frac{\xi}{2}+\left(\frac{k}{B_1}\right)^{\frac{3}{2}}\frac{\sqrt{2}}{n^2}\frac{(\xi^3-3\xi)}{24}\left[\gamma_2^2\left(-\frac{2}{f_1}-\frac{8}{f_2}+\frac{12n}{f_1f_2}+\frac{n}{f_2^2}+\frac{n^2}{f_2^2}\right)+\frac{\gamma_4}{2}\left(-\frac{3}{f_1}-\frac{3}{f_2}+\frac{6n}{f_1f_2}\right)\right]. \tag{2}$$

On the whole this correction is not large. For $n = 20$, $k = 5$ it is

$$-\gamma_2\frac{\xi}{2}0{\cdot}020+\tfrac{1}{24}(\xi^3-3\xi)\,(\gamma_2^2\,0{\cdot}063+\gamma_4\,0{\cdot}010)$$

while for $n = 20$ and $k = 10$ it is

$$-\gamma_2\frac{\xi}{2}0{\cdot}015+\tfrac{1}{24}(\xi^3-3\xi)\,(\gamma_2^2\,0{\cdot}086+\gamma_4\,0{\cdot}013).$$

For the sake of illustration consider the rectangle, where $\gamma_2 = -1{\cdot}2$, $\gamma_4 = 48/7$. The alterations to the tabulated percentage points are as follows

Degrees of freedom	Normal $z_{0{\cdot}05}$	Correction	Normal $z_{0{\cdot}01}$	Correction
$f_1 = 4, f_2 = 15$	0·559	0·016	0·794	0·066
$f_1 = 9, f_2 = 10$	0·553	0·011	0·799	0·072

The corrections are not, on the whole, large and are unlikely to cause serious error in the application of tests of significance. The stability of the 5 % significance level is worthy of note, since it has become apparent in many non-normal investigations. (See, for example, Pearson and Merrington's table 42 in *Biometrika Tables for Statisticians*.)

It will be noticed that while the effect of non-normality with equal groups is to alter the mean and variance of normal z by terms of order n^{-2} the effect of unequal group frequencies is to alter them by terms of order n^{-1}. This means that it is no longer possible to expand about moments of z_N since the Edgeworth expansion requires the assumption that the difference between the second cumulants of z_A and z_N is of order n^{-2}. The simplest thing would seem to be therefore to expand the mean and variance of z_A with unequal group frequencies (z_A^* say) about its own mean and variance, remembering that the cumulants of z_N which can be substituted are given by

$$2\kappa_{1N} = \log f_2 - \log f_1 + \psi(\tfrac{1}{2}f_1) - \psi(\tfrac{1}{2}f_2)$$
$$2^s\kappa_{sN} = \psi^{s-1}(\tfrac{1}{2}f_1) + (-1)^s\psi^{s-1}(\tfrac{1}{2}f_2) \quad (s > 1),$$

where ψ is the digamma function and ψ^{s-1} is its $(s-1)$st differential. For moderate divergencies from the ideal equal group frequency situation, such as envisaged in the expansions for the moments in terms of v given earlier, it is clear from numerical calculation that the effect of these divergencies is to bring the mean and variance of z_A^* back towards those of z_N. The greatest contributions to the mean and to the variance are, in fact,

$$\kappa_{1A^*} = \kappa_{1N} - \gamma_2\left(\frac{k^2A_2v}{4n^3} - \frac{A_1}{2nk}\right),$$
$$\kappa_{2A^*} = \kappa_{2N} + \gamma_2\left(\frac{k^2B_1^2v}{4n^3} - \frac{B_1}{2nk}\right).$$

It is suggested therefore that calculations for the case when the group frequencies are equal, i.e. when $v = 0$, will give a reasonable idea of the magnitude of the error involved if there is non-normality.

The purpose of this present note has been one of method, but E. S. Pearson has pointed out to me that Gayen (1950) also considered the F-test in the one-way classification in some detail. Gayen was concerned in his numerical examples with the effect on the nominal 5% F-level of non-normality and unequal group frequencies and his results are not therefore directly comparable with those obtained here. It is, however, possible to use the same technique to obtain corrections to the tail areas as has been used above to obtain corrections to the percentage point. Thus if $g(\eta)$ is the normal ordinate at η, then for the equal groups case the correction to

$$\int_{-\infty}^{\eta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

will be found to be, to order n^{-1},

$$g(\eta)\left[-\frac{\eta}{2}\frac{(\kappa_{2A}-\kappa_{2N})}{\kappa_{2N}}+\frac{(\eta^3-3\eta)}{24}\frac{(\kappa_{4A}-\kappa_{4N})}{\kappa_{2N}^2}\right]. \tag{3}$$

For reasonable sized k we may write this as

$$g(\eta)\left[\frac{\eta}{2}\frac{\gamma_2}{n}+\frac{(\eta^3-3\eta)}{24}\frac{k^2}{B_1^2}\left(\frac{\gamma_4}{n^2}\left(-\frac{3}{f_1}-\frac{3}{f_2}+\frac{6n}{f_1 f_2}\right)+\frac{2\gamma_2^2}{n^2}\left(-\frac{2}{f_1}-\frac{8}{f_2}+\frac{12n}{f_1 f_2}+\frac{n}{f_2^2}+\frac{n^2}{f_2^2}\right)\right)\right].$$

Somewhat surprisingly, the major part of the correction does not appear to depend on the degrees of freedom, although these are implicit in η; for η is equal to the standardized percentage point of z_N, i.e.

$$\eta = \frac{z_N(\alpha)-\kappa_{1N}}{\sqrt{\kappa_{2N}}}.$$

The comparison with Gayen's work is difficult in that he assumes for the parent population itself an Edgeworth series in which the terms involving γ_4 (of the population) and γ_2^2 are assumed not to exist. It would appear therefore that a comparison should only be made in the region in which the parent population of Gayen is positive definite, i.e. is capable of being a probability distribution. I have therefore chosen a Pearson Type VII for my population with $\gamma_2 = 1$ and $\gamma_4 = 10$. [This is a t-distribution with 10 d.f.] Let the degrees of freedom of the samples be $f_1 = 8, f_2 = 12$. Gayen's table will give the correction to be made to the 0·05 upper level as $-0{\cdot}0030$. Substitution in formula (3) gives the correction as $-0{\cdot}002$ (3). The corrections are therefore of the same order, and (3) can be used to obtain some idea of the error involved for any of the usual significance levels for reasonable k.

I have to thank E. S. Pearson for pointing out errors in the original draft, and H. Scheffé whose remarks on the paucity of arithmetic spurred me into trying to remedy the situation.

REFERENCES

DAVID, F. N. & JOHNSON, N. L. (1951*a*). *Biometrika*, **38**, 43–57.
DAVID, F. N. & JOHNSON, N. L. (1951*b*). *Ann. Math. Statist.* **22**, 382–93.
DAVID, F. N. & JOHNSON, N. L. (1952). *Ann. Math. Statist.* **23**, 594–602.
GAYEN, A. K. (1950). *Biometrika*, **37**, 236–55.
PEARSON, E. S. (1931). *Biometrika*, **23**, 114–33.
WISHART, J. (1947). *Biometrika*, **34**, 170–8.

APPROXIMATION TO THE DISTRIBUTION OF SAMPLE SIZE FOR SEQUENTIAL TESTS

I. TESTS OF SIMPLE HYPOTHESES

By D. H. BHATE

University College London and Delhi School of Economics, New Delhi

1. Introduction

A. Wald (1943) developed a sequential probability ratio test for testing a simple hypothesis against a simple alternative. He gave approximate formulae for the sequential test procedure, its average sample size, and the power of the test. He has also derived the asymptotic form of the distribution of the sample size. The actual distribution of the decisive sample number (D.S.N.) has proved difficult to determine. Kac (1945) gave the distribution of the D.S.N. for the test approximately equivalent to that for comparing two possible values for the mean of a Normal distribution, in terms of the latent roots of an integral equation. Darling & Siegert (1953) have given the distribution of the duration of an absorption process approximating the sequential procedure for testing the mean of a Normal population with known variance. But in practice it is rather difficult to evaluate the actual probability of reaching a decision at a specified value of n. Baker (1950) has studied the distribution of D.S.N. experimentally in a special case. Page (1954*a*) and Albert (1954) have also studied this problem, using Monte Carlo methods and similar techniques. Ramchandran (1954) has developed certain approximations to the distribution of the D.S.N. His method gives satisfactory results for small values, only, of n, because the errors of approximation accumulate with n.

In this paper we shall derive general formulae for upper and lower bounds for the cumulative distribution of the D.S.N., and use them to estimate its approximate distribution.

2. Development of theory

2·1. Let $f(x, \theta)$ be the probability density function of a random variable x, depending on a parameter θ. A sequential procedure, for testing the hypothesis $H_0: \theta = \theta_0$ against the alternative $H_1: \theta = \theta_1$ (H_0 and H_1 may be either simple or composite) may be constructed as follows.

Suppose that $x_1, x_2, \ldots, x_j$ is a sequence of independent observations and that t_j, a function of these x's, is an appropriate test statistic. Then the sequential test procedure is based on a partition of the range of variation of t_j into three mutually exclusive regions C_j, A_j and R_j, using the following rule:

(i) if $t_j \in C_j$, take a further observation;
(ii) if $t_j \in A_j$, accept H_0;
(iii) if $t_j \in R_j$, reject H_0 (i.e. accept H_1).

The regions are chosen in such a way that the probabilities of rejecting H_0 and H_1 when they are true are approximately α and β, respectively.

Following this rule a decision is reached at the ith stage if $t_j \in C_j$ for $j = 1, 2, \ldots, i-1$, and either $t_i \in A_i$ or $t_i \in R_i$. The argument used below to determine the upper and lower bounds

for the distribution of i, the decisive sample number (D.S.N.) may be illustrated with the help of the accompanying diagram, in which j is taken as abscissa and t_j as the ordinate. The boundaries between the regions have been represented by the lines $O'A$, and $O''R$ (not necessarily straight lines).

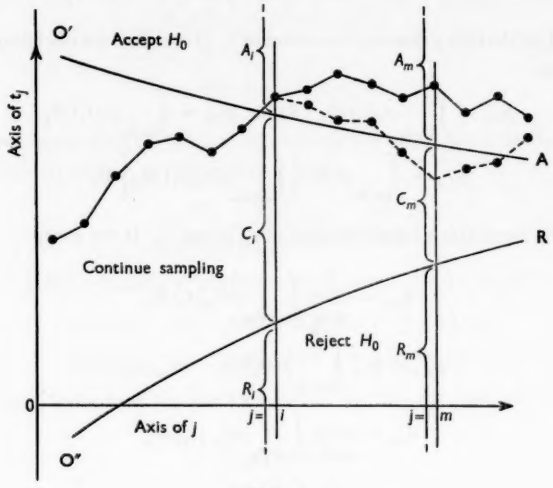


Fig. 1

2·2. Suppose now that although a decision is reached at stage $j = i$, sampling is continued to a fixed stage $j = m$ ($\geqslant i$). In the diagram, two possible sequences of samples are represented. Each track enters the acceptance region at $j = i$ but in one case $t_m \in A_m$ and in the other case $t_m \in C_m$. Similar sequences of samples when the track enters the rejection region at $j = i$ are also possible.

Now define the following probabilities.

(*a*) P_m = probability that a decision is reached at, or before the mth state and $t_m \bar{\in} C_m = P\{t_m \bar{\in} C_m\}$.

(*b*) p_i = probability that a decision is reached at the ith stage (and not before)

$$= P\{t_j \in C_j \quad \text{for} \quad j = 1, 2, \ldots, i-1 \quad \text{and} \quad t_i \in A_i\} + P\{t_j \in C_j \quad \text{for} \quad j = 1, 2, \ldots, i-1 \quad \text{and} \quad t_i \in R_i\} = p_{iA} + p_{iR} \quad \text{(say)}.$$

(*c*) p'_i = probability that a decision is reached at the ith stage (and not before), but that when sampling is continued to the mth stage, $t_m \in C_m$

$$= P\{t_j \in C_j \quad \text{for} \quad j = 1, 2, \ldots, i-1, t_i \in A_i \quad \text{and} \quad t_m \in C_m\} + P\{t_j \in C_j \quad \text{for} \quad j = 1, 2, \ldots, i-1, t_i \in R_i \quad \text{and} \quad t_m \in C_m\}$$

$$= p'_{iA} + p'_{iR} \quad \text{(say)}.$$

The diagram may help to clarify the meaning of these probabilities. The continuous path would be included among those associated with p_{iA} and P_m but not with p'_{iA}; the broken path would be included among those associated with p_{iA} and p'_{iA} but not with P_m.

Then it follows that

$$\sum_{i=1}^{m} p_i = P_m + \sum_{i=1}^{m-1} (p'_{iA} + p'_{iR}). \qquad (2{\cdot}2{\cdot}1)$$

Thus P_m provides a lower limit for the cumulative distribution of the decisive sample number. By considering the upper bound for p'_{iA} and p'_{iR} we shall obtain an upper limit for the cumulative distribution of the D.S.N.

Let
$$p_i(t_i) = p(t_i \mid t_j \in C_j \quad \text{for} \quad j = 1, 2, \ldots, i-1)$$
be the conditional probability density function of t_i, given that no decision has been reached before the ith stage.

Then
$$p_{iA} = \int_{t_i \in A_i} p_i(t_i)\, dt_i \quad \text{and} \quad p_{iR} = \int_{t_i \in R_i} p_i(t_i)\, dt_i.$$

Further
$$p'_{iA} = \int_{t_i \in A_i} p_i(t_i) \left\{ \int_{t_m \in C_m} p(t_m|t_i)\, dt_m \right\} dt_i,$$

where $p(t_m|t_i)$ is the conditional distribution of t_m given t_i. If we write

$$\eta_{im} = \max_{t_i \in A_i} \int_{t_m \in C_m} p(t_m|t_i)\, dt_i, \tag{2·2·2}$$

it follows that
$$p'_{iA} \leqslant \eta_{im} \int_{t_i \in A_i} p_i(t_i)\, dt_i = \eta_{im} p_{iA}.$$

Similarly if
$$\delta_{im} = \max_{t_i \in R_i} \int_{t_m \in C_m} p(t_m \mid t_i)\, dt_i. \tag{2·2·3}$$
$$p'_{iR} \leqslant \delta_{im} p_{iR}.$$

Hence from (2·2·1) we have the bounds

$$P_m \leqslant \sum_{i=1}^{m} p_i \leqslant P_m + \sum_{i=1}^{m-1} \eta_{im}\, p_{iA} + \sum_{i=1}^{m-1} \delta_{im} p_{iR} \tag{2·2·4}$$

The result is valid for any hypothesis H (not only for H_0 and H_1).

Two less sharp, but sometimes more easily calculable, inequalities can be derived from (2·2·4). These are

$$\sum_{i=1}^{m} p_i \leqslant P_m + \sum_{i=1}^{m-1} p_i \eta_{im} + \alpha \max_{i,m} \delta_{im} \quad \text{when } H_0 \text{ is true,} \tag{2·2·5}$$

$$\sum_{i=1}^{m} p_i \leqslant P_m + \beta \max_{i,m} \eta_{im} + \sum_{i=1}^{m-1} p_i \delta_{im} \quad \text{when } H_1 \text{ is true} \tag{2·2·6}$$

We can use (2·2·4) to give useful upper and lower bounds for the cumulative distribution of D.S.N. if η_{im} and δ_{im} can be evaluated, and, being evaluated, are sufficiently small.

Just where, between these limits, the actual value of the cumulative distribution lies, we do not know, but in the examples which follow it appears that a simple average of the upper and the lower bounds gives quite good results.

2·3. Consider now a special case of comparison of two simple hypotheses H_0; H_1 and suppose that we know that, with an obvious notation

$$p_i(H_0) = p_i(H_1).$$

This is so, for example, when H_0 and H_1 specify two values for the mean of a Normal population, and $\alpha = \beta$.

For all decisions to accept H_0, at sample size i

$$p(x_1, \ldots, x_i \mid H_1) < \frac{\beta}{1-\alpha} p(x_1, \ldots, x_i \mid H_0),$$

so
$$p_{iA}(H_1) < \frac{\beta}{1-\alpha} p_{iA}(H_0)$$

and similarly
$$p_{iR}(H_1) > \frac{1-\beta}{\alpha} p_{iR}(H_0).$$

Approximately (i.e. neglecting the amount by which the limits are overshot) the inequalities may be replaced by equalities, giving, when $\alpha = \beta$,

$$\frac{p_{iA}(H_1)}{p_{iA}(H_0)} = \frac{p_{iR}(H_0)}{p_{iR}(H_1)} = \frac{\alpha}{1-\alpha}.$$

Since
$$p_{iA}(H_1) + p_{iR}(H_1) = p_i(H_1) = p_i(H_0) = p_{iA}(H_0) + p_{iR}(H_0),$$

$$\left.\begin{aligned} p_{iA}(H_1) &\doteqdot \alpha p_i \quad \text{and} \quad p_{iA}(H_0) = (1-\alpha) p_i, \\ p_{iR}(H_1) &\doteqdot (1-\alpha) p_i \quad \text{and} \quad p_{iR}(H_0) = \alpha p_i. \end{aligned}\right\} \qquad (2{\cdot}3{\cdot}1)$$

If (2·3·1) be assumed to be true we can proceed to use the inequalities (2·2·4).

Table 1

i	c	i	c	i	c
1	0·0860	6	0·0687	11	0·0686
2	·0778	7	·0686	12	·0686
3	·0710	8	·0686	13	·0686
4	·0693	9	·0686	14	·0686
5	·0688	10	·0686	15	·0686

In order to see how far this assumption is justified we quote figures obtained by quadrature in the course of the calculation of the exact distribution of the D.S.N. for the case

$$\begin{cases} p(x \mid \theta) = \dfrac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}(x-\theta)^2} \\ H_0\colon \theta = \theta_0 = 0, \quad H_1\colon \theta = \theta_1 = 1; \quad \alpha = \beta = 0{\cdot}0758. \end{cases}$$

(The common value 0·0758 for α and β was chosen in Example 2 below, to give simple limits for $t_m = \sum_{j=1}^{m} x_j$.)

The values of the ratio $c = p_{iR}(H_0)/p_i(H_0)$ obtained in this calculation are given in table 1.

Apart from some variation for small i, when some excess of the observed statistic over the boundaries is to be expected, $c = p_{iR}(H_0)/p_i(H_0) \doteqdot 0{\cdot}0686$. Thus the assumption of constancy in the ratio may not be greatly in error, provided the probability of termination for $i = 1, 2, 3, 4$, is fairly small (though the actual value of the ratio is rather less than the nominal value of α).

3. Illustrative examples

Example 1. Let

$$f(x,\theta) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}(x-\theta)^2},$$

$$H_0\colon \theta = \theta_0 = 0, \quad H_1\colon \theta = \theta_1 = 1; \quad \alpha = \beta = 0{\cdot}025.$$

(i) The continuation region at the mth stage, C_m, is defined by the inequalities

$$\frac{0{\cdot}025}{1-0{\cdot}025} < \prod_{i=1}^{m} \frac{f(x_i,1)}{f(x_i,0)} < \frac{1-0{\cdot}025}{0{\cdot}025},$$

or

$$-\log_e 39 < -\frac{1}{2}\sum_{i=1}^{m}[(x_i-1)^2 - x_i^2] < \log_e 39$$

i.e.

$$-3{\cdot}66 + \tfrac{1}{2}m < t_m < 3{\cdot}66 + \tfrac{1}{2}m,$$

where

$$t_m = \sum_{i=1}^{m} x_i.$$

Similarly

(ii) $-3{\cdot}66 + \frac{1}{2}m \geqslant t_m$ for region A_m,

(iii) $3{\cdot}66 + \frac{1}{2}m \leqslant t_m$ for region R_m.

Denoting the probability density function of t_m by $p(t_m)$, where in this case

$$p(t_m) = \frac{1}{\sqrt{(2\pi m)}} \exp\{-\tfrac{1}{2}(t_m - m\theta)^2/m\},$$

we have

$$\eta_{im} = \int_{\frac{1}{2}(m-i)}^{7{\cdot}32+\frac{1}{2}(m-i)} p(t_{m-i})\, dt_{m-i},$$

since this value corresponds to the case when t_i is exactly on the boundary of C_i. Similarly

$$\delta_{im} = \int_{-7{\cdot}32+\frac{1}{2}(m-i)}^{\frac{1}{2}(m-i)} p(t_{m-i})\, dt_{m-i}.$$

The following Table 2 gives, when H_0 is true, the upper and lower bounds for the cumulative distribution of the D.S.N., using (2·2·4), and the average of these bounds. The observed frequencies in the last column are taken from an experimental investigation by Baker (1950).

Comparing the theoretical frequencies Np_i with Baker's sampling results, we find that $\chi^2 = 19{\cdot}0022$ with $\nu = 20$ D.F. (frequencies corresponding to $i = 1$ and 2 were grouped). Thus the somewhat arbitrary assumption that we may approximate to the cumulative sum of the p_i's by taking the mid-point between the upper and lower bounds is not disproved by these results.

Example 2

Let $\alpha = \beta = 0{\cdot}0758$. This case has already been referred to in § 2. The regions C_m, A_m, R_m are now defined by

$$C_m\colon -2{\cdot}50 + \tfrac{1}{2}m < t_m < 2{\cdot}50 + \tfrac{1}{2}m,$$

$$A_m\colon -2{\cdot}50 + \tfrac{1}{2}m \geqslant t_m,$$

$$R_m\colon 2{\cdot}50 + \tfrac{1}{2}m \leqslant t_m.$$

The approximate distribution of the D.S.N. under H_0 can be calculated as in Example 1. In this case, an exact calculation by quadrature was also carried out. At each stage $p_c(t_m)$, the conditional probability density function of t_m, given that no decision had been

reached up to and including the mth stage, was computed by quadrature. The distribution of $t_{m+1} = t_m + x_{m+1}$ was then computed from the formula

$$p_c(t_{m+1}) = \int_{t_m \in C_m} p_c(t_m)\, p(x_{m+1})\, dt_m, \quad \text{where} \quad x_{m+1} = t_{m+1} - t_m.$$

The results of these calculations are shown in Table 3, together with some observed frequencies taken from the unpublished results of experiments carried out by Baker, who kindly put them at my disposal.

Table 2. *Estimation of p_i for Example* 1 *and comparison with Baker's empirical results with* $N = 2002$

Sample no. i	Lower bound L_i	Upper bound U_i	$\frac{1}{2}(U_i + L_i)$	Diff. $= \hat{p}_i$	Expected frequency $N\hat{p}_i$	Observed frequency
1	0·00080	0·00080	0·00080	0·00080	2	0
2	·03055	·03080	·03068	·02988	60	60
3	·10709	·11680	·11194	·08126	163	140
4	·20560	·23914	·22237	·11043	221	217
5	·30451	·36645	·33548	·11311	226	213
6	0·39684	0·48287	0·43985	0·10437	209	219
7	·47944	·58289	·53116	·09131	183	193
8	·55112	·66538	·60825	·07709	154	158
9	·61352	·73289	·67320	·06495	130	125
10	·66583	·78627	·72605	·05285	106	112
11	0·71173	0·82976	0·77074	0·04469	89	99
12	·75438	·86839	·81138	·04064	81	85
13	·78764	·89726	·84245	·03107	62	63
14	·81545	·91855	·86700	·02455	49	54
15	·84090	·93667	·88878	·02178	44	43
16	0·86174	0·95062	0·90618	0·01740	35	41
17	·88059	·96231	·92145	·01527	31	21
18	·89761	·97263	·93512	·01367	27	37
19	·91114	·98000	·94557	·01045	21	26
20	·92334	·98592	·95463	·00906	18	15
21–25	0·93419	—	—	0·02807	56	54
> 25	·96838	—	—	·01730	35	27

Comparing the approximate expected frequencies $N\hat{p}_i$ and the exact expected frequencies Np_i with the observed frequencies we find that $\chi^2 = 18{\cdot}7304$ and $\chi^2 = 13{\cdot}4573$, respectively, for $\nu = 16$. Thus we find that there is agreement between the expected and the observed frequencies. It is clear from this table that the approximate expected frequencies and the exact values of these frequencies agree well. This supports the assumption that we may approximate the cumulative distribution by the mean of the lower and upper bounds.

Table 3. *Values p_i obtained by* (2·2·4) *and by quadrature and their comparison with empirical results with* $N = 698$

Sample no. i	Approximate probabilities $\hat{p}_i$	Exact probabilties p_i	Expected frequencies		Observed frequency
			$N\hat{p}_i$	Np_i	
1	0·02410	0·02410	17	17	17
2	·13126	·13355	92	93	100
3	·16425	·16311	115	114	111
4	·14608	·14459	102	101	100
5	·11702	·11735	82	82	94
6	0·09112	0·09257	64	65	66
7	·07080	·07229	49	50	45
8	·05613	·05628	39	39	37
9	·04326	·04376	30	31	41
10	·03489	·03399	24	24	18
11	0·02796	0·02643	19	18	13
12	·02258	·02052	16	14	12
13	·01844	·01594	13	11	13
14	·01526	·01238	11	9	8
15	·01236	·00963	9	7	7
16	0·01026	0·00747	7	5	3
$\geqslant 17$	·01423	·02604	10	18	13

Example 3

Let x be a Normal variate with zero mean and variance σ^2.

$$H_0: \sigma = \sigma_0 = 1, \quad H_1: \sigma = \sigma_1 = 3; \quad \alpha = \beta = 0{\cdot}01.$$

The test procedure is then based on the following regions:

$$C_m: -10{\cdot}350 + 2{\cdot}475m < \sum_{j=1}^{m} x_j^2 < 10{\cdot}350 + 2{\cdot}475m,$$

$$A_m: -10{\cdot}350 + 2{\cdot}475m \geqslant \sum_{j=1}^{m} x_j^2,$$

$$R_m: 10{\cdot}350 + 2{\cdot}475m \leqslant \sum_{j=1}^{m} x_j^2.$$

In this case we cannot assume that p_{iA}/p_i is approximately constant, but will use one of the bounds given by (2·2·5). We shall calculate the approximate distribution of the D.S.N. under the null hypothesis, and will use the formula

$$P_m \leqslant \sum_{i=1}^{m} p_i \leqslant P_m + \sum_{i=1}^{m-1} p_i \eta_{im} + \alpha \max_{i,m} \delta_{im}.$$

We find that $\max_{i,m} \delta_{im}$ is very nearly equal to 1 for a number of combinations of i and m and therefore we can simplify our calculations by calculating the bounds from the formula

$$P_m \leqslant \sum_{i=1}^{m} p_i \leqslant P_m + \sum_{i=1}^{m-1} p_i \eta_{im} + \alpha,$$

where

$$\eta_{im} = \int_{2\cdot475(m-i)}^{20\cdot700+2\cdot475(m-i)} p(\chi^2_{m-i})\, d\chi^2_{m-i}$$

$$\left(\text{since } \sum_{j=i+1}^{m} x_j^2 \text{ is distributed as } \chi^2_{m-i}\sigma^2\right).$$

The following table gives the approximate D.S.N. distribution, together with the upper and the lower bounds for the cumulative D.S.N. distribution from which the approximate distribution was obtained by averaging.

Table 4. *Estimation of p_i for Example* 3

Sample no. i	Lower bound L_i	Upper bound U_i	Mean of the bounds $\frac{1}{2}(L_i+U_i)$	Approximate probabilities $\hat{p}_i$
5	0·13722	0·13722	0·13722	0·13722
6	·39140	·41729	·40434	·26712
7	·56689	·61939	·59314	·18880
8	·69340	·75623	·72481	·13167
9	·77997	·84338	·81167	·08686
10	0·84448	0·90972	0·87360	0·06193
11	·88878	·94038	·91458	·04098
12	·91951	·96396	·94178	·02715
13	·94342	·98146	·96242	·02069
14	·95838	·99105	·97472	·01230
$\geqslant 15$	0·96991	—	1·00000	0·02528

The region A_m does not exist for values of m less than 5 and therefore decision to accept H_0 is not possible for these values of m; hence the above calculations are for values of $m \geqslant 5$. The probability p_5 contains, in addition to the probability of reaching a decision at $m = 5$, the probability of reaching a decision to reject H_0 at $m = 1, 2, 3$ or 4 which is almost negligible.

4. Bounds for the average sample number (A.S.N.) of a sequential test

If we write, abbreviating (2·2·4),

$$P_m \leqslant \sum_{i=1}^{m} p_i \leqslant P'_m,$$

then it is easy to show that, for sufficiently large N,

$$N+1-\sum_{m=1}^{N} P'_m \leqslant \sum_{j=1}^{N} jp_j \leqslant N+1-\sum_{m=1}^{N} P_m. \tag{4·1}$$

As N tends to infinity the central term in these inequalities tends to the A.S.N. of the test. In Table 5 the limits for the A.S.N. so obtained are compared with the limits given by Wald for the problems stated in Examples 1 and 2.

The experimental value 8·8 of the A.S.N. was taken from Baker (1950) and is based on the results of 2002 tests. The other experimental value 5·65 is based on the unpublished results of the 698 tests carried out by Baker.

From this table we can see that the lower bound, given by (4·1), is sharper than that given by Wald, but that Wald gives the sharper upper bound. Hence, the lower bound in (4·1), and Wald's upper bound, give narrower limits than either system alone.

Table 5. *Comparison of upper and the lower bounds for the* A.S.N. *obtained by using Wald's formula with those obtained using equation* (4·1·1). *Case* H_0: $\theta = 0$; H_1: $\theta = 1$

Risks ...	$\alpha = \beta = 0{\cdot}025$			$\alpha = \beta = 0{\cdot}0758$			
Result obtained by ...	Wald	Equation (4·1·1)	Expt.	Wald	Equation (4·1·1)	Expt.	Exact
Lower bound	6·9	7·61	—	4·20	4·94	—	—
Approx. A.S.N.	7·2	8·62	8·8	4·24	5·81	5·65	5·87
Upper bound	8·9	9·64	—	6·10	7·20	—	—

5. Conclusion

Many of the results in this paper are valid for tests of composite hypothesis also. It is hoped to publish the results of similar investigations for composite hypothesis in the near future.

The author wishes to thank Dr H. O. Hartley and Dr N. L. Johnson for their valuable suggestions and guidance, while the author was working in the Department of Statistics, University College London, and also the referee for suggestions and criticisms.

REFERENCES

Albert, G. E. (1954). On the computation of sampling characteristics of a general class of sequential decision problems. *Ann. Math. Statist.* **25**, 340–56.

Baker, A. G. (1950). Properties of some tests in sequential analysis. *Biometrika*, **37**, 334–46.

Bhate, D. H. (1958). A note on the distribution of the successive sums of samples from an exponential population. *Bull. Calcutta Statist. Ass.* **8**, 13–19.

Darling, D. A. & Siegert, A. T. (1953). First passage problem for a continuous Markoff process. *Ann. Math. Statist.* **24**, 624–39.

Kac, M. (1945). Random walk in the process of absorbing barriers. *Ann. Math. Statist.* **14**, 62–7.

Page, E. S. (1954*a*). Monte Carlo solutions of some integral equations. *Proc. Camb. Phil. Soc.* **50**, 414–26.

Page, E. S. (1954*b*). An improvement of Wald's approximations for some properties of sequential tests. *J.R. Statist. Soc.* Ser. B, **16**, 136–9.

Ramchandran, G. (1954). Some contributions to sequential analysis. Ph.D. thesis, University of Madras.

Wald, A. (1943). *Sequential Analysis of Statistical Data: Theory.* Report submitted to Applied Mathematics Panel; National Defence Research Committee.

Wald, A. (1947). *Sequential Analysis.* New York: John Wiley and Sons.

TOURNAMENTS AND PAIRED COMPARISONS*

By H. A. DAVID
Virginia Polytechnic Institute and University of Melbourne

1. Introduction and summary

In an experiment involving paired comparisons t 'stimuli' such as flavours or colours are presented in pairs to a panel of n judges. The basic experimental unit is the comparison of two stimuli by a single judge who must state which one he prefers. Such an individual paired comparison may be regarded as a game between two contestants—the stimuli—while the experiment as a whole is analogous to a tournament of t players. If all the possible $\frac{1}{2}t(t-1)$ comparisons are made by each of the n judges, this corresponds to a Round Robin tournament repeated n times. It will be convenient in this paper to use the language of tournaments rather than that of paired comparisons.

The analogy between a paired comparison experiment and a Round Robin tournament, first pointed out by Kendall (1955), suggests the investigation of another well-known way of running competitions, the Knock-out tournament. It is perhaps surprising that such tournaments do not appear to have been studied at all in the statistical literature, although this may be due to their inherent lack of balance†. However, while this objection is valid if a ranking of all players is desired, it loses much of its force when the aim is to pick the best player. In § 2 some properties of a simple (i.e. unrepeated) Knock-out tournament are investigated; in particular, an expression is found for the probability with which the best player wins such a tournament, various assumptions about the strengths of the players being made.

Round Robin tournaments (in effect) have received considerable attention both from statisticians and psychometricians. One problem is the enumeration of the possible outcomes, expressed as scores, of a tournament of a given size and the determination of the frequencies of these outcomes. The matter has been investigated by Bradley & Terry (1952) and Bradley (1954) and is further considered in § 3 and Table 1. Bradley & Terry have also proposed a model to represent the strengths $\pi_i (i = 1, 2, \ldots, t; \Sigma\pi_i = 1)$ of the players which allows the π_i to be estimated from the players' scores. A different model had been suggested earlier by Thurstone (1927). Tests of the over-all equality of scores may be based on these methods or on combinatorial arguments leading to a test criterion involving the sum of squares of the scores (Durbin, 1951). For a simple Round Robin tournament the latter approach is equivalent to the 'circular triad test' of Kendall & Babington Smith (1940).

However, these approaches do not provide an answer to such questions as what constitutes a significant difference in the scores of two specified players. To deal with problems of this type, the joint probability distribution of the scores of any $s(\leqslant t)$ players in a *simple* Round Robin tournament is found in § 3·1 on the hypothesis, H_0, that the players are all of equal strength. This result is used to obtain tests of significance based on what are essentially upper percentage points of (i) the score of a specified player, (ii) the difference in the scores of two specified players, and (iii) the top score in the tournament.

Some aspects of repeated tournaments are discussed in § 4.

* Research supported, in part, by the Office of Ordnance Research, U.S. Army.

† [See, however, a paper by Rita J. Maurice (*Biometrika* (1958), **45**, 581–6) published since this paper was written. Ed.].

2. Knock-out tournaments

We suppose that the number t of players has been reduced, by preliminary play-offs or otherwise, to a power of 2, say $t = 2^p$. The tournament will then be completed in p rounds. If all contestants are of equal strength, each has probability $t^{-1} = 2^{-p}$ of winning the tournament; or, more briefly, each has *winning probability* 2^{-p}. Likewise, if a particular player beats any one opponent with probability π, he will win the tournament with probability π^p.

In more complicated cases the winning probability of a specified player will evidently depend on the initial tournament draw, which we take to be strictly random. Let us call the t contestants $C_1, C_2, \ldots, C_t$ and write

$$C_i \xrightarrow[\pi]{} C_j \quad (i, j = 1, 2, \ldots, t;\ i \neq j),$$

to express that C_i defeats C_j with probability π. Let $P_i^{(r)}$ $(r = 1, 2, \ldots, p)$ denote the probability that C_i reaches the rth round of the tournament *and* wins his game in this round, so that $P_i^{(p)}$ is his probability of winning the tournament. Finally, let $M_{ij}^{(r)}$ be the probability that C_i meets C_j in the rth round and M_{ij} be the probability of their meeting in the course of the tournament. Clearly

$$M_{ij} = \sum_{r=1}^{p} M_{ij}^{(r)}. \tag{2·1}$$

Suppose now that

$$C_1 \xrightarrow[\pi_1]{} C_3, C_4, \ldots, C_t,$$
$$C_2 \xrightarrow[\pi_2]{} C_3, C_4, \ldots, C_t,$$
$$C_1 \xrightarrow[\pi_{12}]{} C_2.$$

Then C_1's probability of winning the tournament is given by

$$P_1^{(p)} = M_{12}\pi_{12}\pi_1^{p-1} + (1 - M_{12})\pi_1^p. \tag{2·2}$$

If the draw is known, M_{12} and hence $P_1^{(p)}$ can be evaluated. Instead of considering this, we shall obtain expressions for $\overline{M}_{12}$ and $\overline{P}_1^{(p)}$, the values of M_{12} and $P_1^{(p)}$ *before* the draw. These prior probabilities will still be linked by (2·2). We have

$$\overline{M}_{12}^{(1)} = 1/(t-1).$$

If C_1 and C_2 are to meet in the second round, in which there will be $\frac{1}{2}t$ contestants, they must both be successful in the first round, so that

$$\begin{aligned}\overline{M}_{12}^{(2)} &= (1 - \overline{M}_{12}^{(1)})\,\pi_1\pi_2/(\tfrac{1}{2}t - 1)\\ &= 2\pi_1\pi_2/(t-1).\end{aligned}$$

Continuing the argument, we find

$$\overline{M}_{12}^{(r)} = (2\pi_1\pi_2)^{r-1}/(t-1),$$

and hence from (2·1)

$$\overline{M}_{12} = \sum_{r=1}^{p} (2\pi_1\pi_2)^{r-1}/(t-1).$$

It is interesting to note that the $\overline{M}_{ij}^{(r)}$ depend on π_1 and π_2 only through the product $\pi_1\pi_2$. We see also that the probability of a final between C_1 and C_2 is

$$\overline{M}_{12}^{(p)} = (2\pi_1\pi_2)^{p-1}/(t-1).$$

If for the moment we imagine C_1 and C_2 to be 'seeded' players placed in opposite halves of an otherwise random draw, this probability would become

$$(\pi_1\pi_2)^{p-1} \doteq 2\overline{M}_{12}^{(p)}.$$

We turn now to the general case where C_i defeats C_j with probability π_{ij}, the only restriction on the π_{ij} being $\pi_{ji} = 1 - \pi_{ij}$ (due to the absence of ties). It should be noted that the possibility

$$\pi_{ij} > \tfrac{1}{2}, \quad \pi_{jk} > \tfrac{1}{2}, \quad \pi_{ki} > \tfrac{1}{2}$$

is not excluded (compare Kendall & Babington Smith, 1940). This situation can occur when the scale of playing strength is not linear, in which case we define the best player as the one with the largest expected score. The models of Thurstone and Bradley & Terry assume a linear scale corresponding to the choice of suitable expressions for the π_{ij}. As we are not concerned with estimation, they need no separate treatment here.

To illustrate the general case take $t = 8$ and suppose (without loss of generality) that $C_1, \ldots, C_8$ have been drawn in that order. Then we have, for example,

$$P_1^{(1)} = \pi_{12}, P_2^{(1)} = \pi_{21}, P_3^{(1)} = \pi_{34}, \ldots, \tag{2·3}$$

$$P_1^{(2)} = P_1^{(1)}(\pi_{13}P_3^{(1)} + \pi_{14}P_4^{(1)}) = \pi_{12}(\pi_{13}\pi_{34} + \pi_{14}\pi_{43}), \tag{2·4}$$

$$P_1^{(3)} = P_1^{(2)}(\pi_{15}P_5^{(2)} + \pi_{16}P_6^{(2)} + \pi_{17}P_7^{(2)} + \pi_{18}P_8^{(2)}). \tag{2·5}$$

The corresponding probabilities before the draw can be obtained by averaging over all possible draws. Let $P_1^{(2)}(i,j,k)$ denote the joint probability that C_1 beats C_i in the first round and the winner of $C_j v. C_k$ in the second round; then $P_1^{(2)}(2,3,4)$ is the expression of (2·4). Likewise, we may (for $t = 8$) write $P_1^{(3)}$ of (2·5) more fully as $P_1^{(3)}(2,3,4)$. It follows that

$$\bar{P}_1^{(1)} = \frac{1}{7}\sum_{i=2}^{8}\pi_{1i}, \tag{2·6}$$

$$\bar{P}_1^{(2)} = \tfrac{1}{105}\sum_{\substack{i \\ 1+i+j+k}}\sum_{j}\sum_{<k} P_1^{(2)}(i,j,k), \tag{2·7}$$

$$\bar{P}_1^{(3)} = \tfrac{1}{105}\sum_{\substack{i \\ 1+i+j+k}}\sum_{j}\sum_{<k} P_1^{(3)}(i,j,k).$$

For $t > 8$ the number of terms on the right-hand side may become formidable. Thus $\bar{P}_1^{(r)}$ is the mean of

$$(t-1)\binom{t-2}{2}\binom{t-4}{4}\cdots\binom{t-2^{r-1}}{2^{r-1}} \text{ terms.}$$

Example. In a tournament of 8 players, let C_i's 'strength' in his encounter with C_j be represented by

$$x_{ij} = \xi_i + v_{ij},$$

where ξ_i is C_i's expected strength and the error variate v_{ij} is $N(0, \frac{1}{2}\sigma^2)$, the $56\,v_{ij}$'s being mutually independent (compare Mosteller, 1951).

Then

$$\pi_{ij} = \Pr(\xi_j + v_{ji} < \xi_i + v_{ij}) = \int_{-\infty}^{(\xi_i - \xi_j)/\sigma} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}u^2}\,du.$$

We now take ξ_i to be the expected value of the ith largest unit normal deviate in samples of 8, so that the players are arranged in decreasing order of strength with C_1 the strongest. The ξ_i have been tabulated by Teichroew (1956). For $\sigma = 1$ the following array of π_{ij}'s results:

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
C_1	—	0·7161	0·8291	0·8982	0·9425	0·9710	0·9886	0·9978
C_2	0·2839	—	0·6478	0·7579	0·8425	0·9074	0·9558	0·9886
C_3	0·1709	0·3522	—	0·6256	0·7341	0·8278	0·9074	0·9710
C_4	0·1018	0·2421	0·3744	—	0·6198	0·7341	0·8425	0·9425
C_5	0·0575	0·1575	0·2659	0·3802	—	0·6256	0·7579	0·8982
C_6	0·0290	0·0926	0·1722	0·2659	0·3744	—	0·6478	0·8291
C_7	0·0114	0·0442	0·0926	0·1575	0·2421	0·3522	—	0·7161
C_8	0·0022	0·0114	0·0290	0·0575	0·1018	0·1709	0·2839	—

It will be seen that due to the symmetry of the normal distribution

$$\pi_{ij} = \pi_{9-j,9-i}.$$

Applying (2·6) and (2·7) we obtain

$$\bar{P}_i^{(1)} = 0{\cdot}906,\ 0{\cdot}769,\ 0{\cdot}656,\ 0{\cdot}551,\ 0{\cdot}449,\ 0{\cdot}344,\ 0{\cdot}231,\ 0{\cdot}094;$$
$$\bar{P}_i^{(2)} = 0{\cdot}781,\ 0{\cdot}494,\ 0{\cdot}319,\ 0{\cdot}200,\ 0{\cdot}117,\ 0{\cdot}060,\ 0{\cdot}024,\ 0{\cdot}004.$$

It is interesting to compare these probabilities before the draw with those in a draw in which considerable seeding has been used, the first round games being $C_1 v. C_8$, $C_5 v. C_4$, $C_3 v. C_6$, $C_7 v. C_2$. From (2·3) to (2·5) we find

$$P_i^{(1)} = 0{\cdot}998,\ 0{\cdot}956,\ 0{\cdot}828,\ 0{\cdot}620,\ 0{\cdot}380,\ 0{\cdot}172,\ 0{\cdot}044,\ 0{\cdot}002;$$
$$P_i^{(2)} = 0{\cdot}913,\ 0{\cdot}662,\ 0{\cdot}312,\ 0{\cdot}064,\ 0{\cdot}023,\ 0{\cdot}020,\ 0{\cdot}006,\ 0{\cdot}002;$$
$$P_i^{(3)} = 0{\cdot}692,\ 0{\cdot}216,\ 0{\cdot}066,\ 0{\cdot}019,\ 0{\cdot}005,\ 0{\cdot}001,\ 0{\cdot}000,\ 0{\cdot}000.$$

3. Round Robin tournaments

The results of a Round Robin tournament can be presented fully in the familiar two-way table of 1's and 0's, with the scores (i.e. number of wins) of each player listed in a column of totals. As a summary of the scores, we may write $(a_1 a_2 \dots a_t)$, meaning thereby that C_1 has won a_1 games, C_2 has won a_2 games, and so on; clearly $\Sigma a_i = \frac{1}{2}t(t-1)$. If the a's differ greatly, a real difference between at least some of the players is indicated. Thus a Round Robin tournament, unlike a Knock-out tournament—or a ranking—provides the possibility of a test of the null hypothesis, H_0, that all players are of equal strength even when there has been no replication. From this point of view the identity of the players is of no consequence, and we may replace $(a_1 a_2 \dots a_t)$ by the partition $[x_1^{r_1} x_2^{r_2} \dots x_m^{r_m}]$, where the x's are simply a re-arrangement of the a's in decreasing order of magnitude, and r_1 is the number of occurrences of the largest score x_1, etc.* It follows that

$$\sum_{u=1}^{m} r_u = t, \quad \sum_{u=1}^{m} r_u x_u = \tfrac{1}{2}t(t-1).$$

One test of H_0 is the 'circular triad test' due to Kendall & Babington Smith (1940); another has been given by Bradley & Terry (1952) and is based on their model referred to earlier. In both cases the test-criterion is a symmetric function of the scores, and its null distribution is easily deduced from a knowledge of the different permissible partitions of $\frac{1}{2}t(t-1)$ together with their frequencies of occurrence. Such basic tables have been constructed by Bradley & Terry (1952), and Bradley (1954) for $t \leqslant 5$. These authors also show how results for several complete repetitions may be built up from those for $n = 1$ and their tables cover many such cases. In Table 1 we have extended their results for $n = 1$ to $t \leqslant 8$, using a score of 1 for a win and 0 for a loss, rather than 1 for a win and 2 for a loss.

A possible approach to the enumeration problem involved is to consider the generating function

$$G(t) = \prod_{i<j} (b_i + b_j), \tag{3·1}$$

a product of $\frac{1}{2}t(t-1)$ terms corresponding to the games played. The expansion of $G(t)$ contains $2^{\frac{1}{2}t(t-1)}$ terms, the exponents in each of which correspond to a possible set of scores. For example,

$$G(3) = (b_1 + b_2)(b_1 + b_3)(b_2 + b_3) = \Sigma b_1^2 b_2 + 2 b_1 b_2 b_3, \tag{3·2}$$

* We depart slightly from standard notation in showing zeros explicitly.

Table 1. *Partitions of scores and their frequencies in a Round Robin tournament of t players*

$t = 3$	
$[210]$	6
1^3	2
	–
Total	8

$t = 4$	
$[3210]$	24
31^3	8
2^30	8
2^21^2	24
	—
Total	64

$t = 5$	
$[43210]$	120
431^3	40
42^30	40
42^21^2	120
3^310	40
3^22^20	120
3^21^2	240
32^31	280
2^5	24
	—
Total	1,024

$t = 6$	
$[543210]$	720
5431^3	240
542^30	240
542^21^2	720
53^310	240
53^22^20	720
53^21^2	1,440
532^31	1,680
52^5	144
4^3210	240
4^31^3	80
4^23^210	720
4^232^20	1,440
4^321^2	2,880
4^22^31	1,680
43^20	1,680
43^21^2	1,680
43^22^21	8,640
432^4	2,400

$t = 6$ (*cont.*)	
$[3^32^3]$	2,640
3^421	2,400
3^50	144
	—
Total	32,768

$t = 7$	
$[6543210]$	5,040
65431^3	1,680
6542^30	1,680
6542^21^2	5,040
653^310	1,680
653^22^20	5,040
653^21^2	10,080
6532^31	11,760
652^5	1,008
64^3210	1,680
64^31^3	560
64^23^210	5,040
64^232^20	10,080
64^321^2	20,160
64^22^31	11,760
643^20	11,760
643^21^2	11,760
643^22^21	60,480
6432^4	16,800
63^50	1,008
63^421	16,800
63^23	18,480
5^3210	1,680
5^31^3	560
5^22^20	560
$5^22^21^2$	1,680
5^24^2210	5,040
$5^24^21^3$	1,680
5^243^210	10,080
5^2432^20	20,160
5^24321^2	40,320
5^242^31	23,520
5^23^220	11,760
$5^23^21^2$	11,760
$5^23^22^21^2$	60,480
5^232^4	16,800
54^3210	11,760
54^32^20	11,760

$t = 7$ (*cont.*)	
$[54^321^2]$	23,520
54^23^220	60,480
$54^23^21^2$	60,480
54^232^31	206,640
54^32^4	28,560
543^40	16,800
543^321	223,440
543^22^3	183,120
53^51	17,808
53^42^2	72,240
4^510	1,008
4^4320	16,800
4^431^2	16,800
4^42^21	28,560
4^33^30	18,480
4^33^221	183,120
4^332^3	99,680
4^23^41	72,240
$4^23^22^2$	233,520
43^52	72,240
3^7	2,640
	—
Total	2,097,152

$t = 8$	
$[76543210]$	40,320
765431^3	13,440
76542^30	13,440
76542^21^2	40,320
7653^310	13,440
7653^22^20	40,320
7653^21^2	80,640
76532^31	94,080
7652^5	8,064
764^3210	13,440
764^31^3	4,480
764^23^210	40,320
764^232^20	80,640
764^321^2	161,280
764^22^31	94,080
7643^20	94,080
7643^21^2	94,080
7643^22^21	483,840
76432^4	134,400
763^50	8,064

$t = 8$ (*cont.*)	
$[763^421]$	134,400
763^23^2	147,840
75^33210	13,440
75^331^3	4,480
75^32^30	4,480
$75^32^21^2$	13,440
75^24^2210	40,320
$75^24^21^3$	13,440
75^243^210	80,640
75^2432^20	161,280
75^24321^2	322,560
75^242^31	188,160
75^23^220	94,080
$75^23^21^2$	94,080
$75^23^22^21$	483,840
75^232^4	134,400
754^3310	94,080
754^32^20	94,080
754^321^2	188,160
754^23^220	483,840
$754^23^21^2$	483,840
754^232^31	1,653,120
754^22^4	228,480
7543^40	134,400
7543^321	1,787,520
7543^22^3	1,464,960
753^51	142,464
753^42^2	577,920
74^510	8,064
74^4320	134,400
74^431^2	134,400
74^42^21	228,480
74^33^30	147,840
74^33^221	1,464,960
74^332^3	797,440
74^23^41	577,920
$74^23^22^2$	1,868,160
743^52	577,920
73^7	21,120
6^43210	13,440
6^4321^3	4,480
6^42^30	4,480
$6^42^21^2$	13,440
6^33^210	4,480
$6^33^22^20$	13,440

Table 1 (*cont.*)

$t = 8$ (*cont.*)		$t = 8$ (*cont.*)		$t = 8$ (*cont.*)		$t = 8$ (*cont.*)	
$[6^3 3^2 2 1^2]$	26,880	$[6^2 4 3^2 2^3]$	1,464,960	$[6543^2 1]$	3,790,080	$[5^3 43^2 21]$	7,889,280
$6^3 32^3 1$	31,360	$6^2 3^5 1$	142,464	$6543^2 2^2$	12,230,400	$5^3 432^3$	4,282,880
$6^3 2^5$	2,688	$6^2 3^4 2^2$	577,920	$653^5 2$	1,876,224	$5^3 3^4 1$	1,025,920
$6^2 5^2 3210$	40,320	$65^3 4210$	94,080			$5^3 3^3 2^2$	3,306,240
$6^2 5^2 31^3$	13,440			$64^5 20$	142,464	$5^2 4^2 20$	577,920
		$65^3 41^3$	31,360	$64^5 1^2$	142,464	$5^2 4^2 1^2$	577,920
$6^2 5^2 2^2 0$	13,440	$65^3 3^2 10$	94,080	$64^4 3^2 0$	577,920	$5^2 4^2 3^2 0$	1,868,160
$6^2 5^2 2^2 1^2$	40,320	$65^3 32^2 0$	188,160	$64^4 321$	3,790,080		
$6^2 54^2 210$	80,640	$65^3 321^2$	376,320	$64^4 2^3$	1,025,920	$5^2 4^3 321$	12,230,400
$6^2 54^2 1^3$	26,880	$65^3 2^3 1$	219,520			$5^2 4^3 2^3$	3,306,240
$6^2 543^2 10$	161,280			$64^3 3^3 1$	3,924,480	$5^2 4^2 3^3 1$	9,461,760
		$65^2 4^2 310$	483,840	$64^3 3^2 2^2$	9,461,760	$5^2 4^2 3^2 2^2$	22,780,800
$6^2 5432^2 0$	322,560	$65^2 4^2 2^2 0$	483,840	$64^2 3^4 2$	7,203,840	$5^2 43^4 2$	11,518,080
$6^2 54321^2$	645,120	$65^2 4^2 21^2$	967,680	643^6	725,760		
$6^2 542^3 1$	376,320	$65^2 43^2 20$	1,653,120	$5^5 210$	8,064	$5^2 3^6$	577,920
$6^2 53^3 20$	188,160	$65^2 43^2 1^2$	1,653,120			$54^5 30$	577,920
$6^2 53^3 1^2$	188,160			$5^5 1^3$	2,688	$54^5 21$	1,876,224
		$65^2 432^2 1$	5,644,800	$5^4 4310$	134,400	$54^4 3^2 1$	7,203,840
$6^2 53^2 2^2 1$	967,680	$65^2 42^4$	779,520	$5^4 42^2 0$	134,400	$54^4 32^2$	11,518,080
$6^2 532^4$	268,800	$65^2 3^4 0$	228,480	$5^4 421^2$	268,800		
$6^2 4^3 310$	94,080	$65^2 3^3 21$	3,037,440	$5^4 3^2 20$	228,480	$54^3 3^3 2$	23,157,120
$6^2 4^3 2^2 0$	94,080	$65^2 3^2 2^3$	2,486,400			$54^3 2^5$	5,201,280
$6^2 4^3 21^2$	188,160			$5^4 3^2 1^2$	228,480	$4^7 0$	21,120
		$654^4 10$	134,400	$5^4 32^2 1$	779,520	$4^6 31$	725,760
$6^2 4^2 3^2 20$	483,840	$654^3 320$	1,787,520	$5^4 2^4$	107,520	$4^6 2^2$	577,920
$6^2 4^2 3^2 1^2$	483,840	$654^3 31^2$	1,787,520	$5^3 4^2 10$	147,840		
$6^2 4^2 32^2 1$	1,653,120	$654^3 2^2 1$	3,037,440	$5^3 4^2 320$	1,464,960	$4^5 3^2 2$	5,201,280
$6^2 4^2 2^4$	228,480	$654^2 3^3 0$	1,464,960			$4^4 3^4$	3,230,080
$6^2 43^4 0$	134,400			$5^3 4^2 31^2$	1,464,960		
		$654^2 3^2 21$	14,515,200	$5^3 4^2 2^2 1$	2,486,400	Total	268,435,456
$6^2 43^3 21$	1,787,520	$654^2 32^3$	7,889,280	$5^3 43^3 0$	797,440		

the sum being over 6 terms. This identity may be interpreted as follows. In a tournament of 3 players there are 6 outcomes of type [210] where one player wins both his games, a second wins one game and the third wins none; and 2 outcomes of type $[1^3]$ where each player wins one game. $G(t)$ is symmetric in the a's and may be expanded as in (3·2) as a sum of monomial symmetric functions. Table 1 has been constructed essentially by noting that

$$G(t+1) = G(t) \prod_{i=1}^{t} (b_{t+1} + b_i) = G(t)[b_{t+1}^t + (1)\, b_{t+1}^{t-1} + (1^2)\, b_{t+1}^{t-2} + \ldots + (1^t)],$$

where $$(1^s) = \Sigma b_1 b_2 \ldots b_s, \quad (s = 1, 2, \ldots, t).$$

Three generalizations of (3·1) are perhaps worth pointing out, although they do not necessarily provide an easy solution to the problems involved.

(*a*) If C_i and C_j play n_{ij} games instead of merely one, the appropriate generating function is

$$\prod_{i<j} (b_i + b_j)^{n_{ij}}.$$

(*b*) If $C_i \xrightarrow[\pi_{ij}]{} C_j$ for all $i \neq j$ then $G(t)$ can be replaced by

$$\prod_{i<j} (b_i \pi_{ij} + b_j \pi_{ji})$$

to generate the *probabilities* of the various partitions.

(*c*) Suppose that instead of being two-point (1 and 0) the scoring scale is $(r+1)$-point $\{1, (r-1)/r, \ldots, (1/r), 0\}$, with the total score in any one game still unity. Then

$$\prod_{i<j} (b_i + b_i^{(r-1)/r} b_j^{1/r} + b_i^{(r-2)/r} b_j^{2/r} + \ldots + b_j)$$

generates all possible outcomes. This includes the case where games may end in either a win, a loss or a tie ($r = 2$).

3·1. *Significance tests*

As already mentioned, several methods are available for testing the null hypothesis H_0 that all t players in a Round Robin tournament are of equal strength. These tests are exact only for small values of t, but approximations for large t have also been given. However, they are overall tests and do not provide an answer to the more detailed questions we may wish to ask. We consider the following questions:

(i) What score must a *specified* player reach to be regarded as significantly better than average?

(ii) What constitutes a significant difference in scores between two specified players?

(iii) On the basis only of the scores obtained, what constitutes a significantly high score?

In order to deal with these problems the joint distribution of scores, on H_0, of $s\,(\leqslant t)$ players is required. C_i's score clearly has the binomial distribution

$$f(a_i) = \binom{t-1}{a_i} \frac{1}{2^{t-1}}.$$

To find the joint distribution of a_i and a_j, we note that these scores arise if C_i defeats C_j and wins $a_i - 1$ of his other $t-2$ games, while C_j wins a_j; and similarly if C_j defeats C_i. Thus

$$f(a_i, a_j) = \frac{1}{2^{2t-3}} \left[\binom{t-2}{a_i - 1} \binom{t-2}{a_j} + \binom{t-2}{a_i} \binom{t-2}{a_j - 1} \right]. \tag{3·3}$$

The argument may be extended to give the joint distribution of s players $C_i, C_j, \ldots, C_k$ by a consideration of the $2^{\frac{1}{2}s(s-1)}$ outcomes of the $\frac{1}{2}s(s-1)$ games between these s players. If their scores in this sub-tournament are $a_i', a_j', \ldots, a_k'$, their final scores will be $a_i, a_j, \ldots, a_k$ with probability

$$f(a_i, a_j, \ldots, a_k \mid a_i', a_j', \ldots, a_k') = \frac{1}{2^{s(t-s)}} \binom{t-s}{a_i - a_i'} \binom{t-s}{a_j - a_j'} \cdots \binom{t-s}{a_k - a_k'},$$

provided we use the convention

$$\binom{n}{r} = 0 \quad \text{for } r < 0 \quad \text{and} \quad r > n. \tag{3·4}$$

It follows that

$$f(a_i, a_j, \ldots, a_k) = \frac{1}{2^{\frac{1}{2}s(2t-s-1)}} \Sigma \prod_{l=i,j,\ldots,k} \binom{t-s}{a_l - a_l'}, \tag{3·5}$$

where the sum is, in effect, over all outcomes of the sub-tournament compatible with the final scores.

We turn now to questions (i)–(iii) above.

(*i*) *Test of the score of a specified player*

Let C_1 be the player in whose performance we are interested. Since, on H_0, a_1 is a binomial variate with parameters $t-1$, $\frac{1}{2}$, it is easy to test H_0 against the alternative that C_1 is above average strength (one-sided binomial test) or different from the average strength of the other players (two-sided binomial test). For example, at the 5 % level of significance the

following are the smallest values of a_1 which are significantly higher than the expected value $\frac{1}{2}(t-1)$:

t	a_1
$\leqslant 5$	no significantly high score possible
6–8	$t-1$
9–11	$t-2$
> 11	smallest integer $> \frac{1}{2}[t+1{\cdot}645\sqrt{(t-1)}]$.

The large-sample result follows from the normal approximation to the binomial, with continuity correction.

Next, suppose that the players are not of equal strength, but that it is desired to test whether C_1's average probability of success $\pi_{1.}$ equals $\frac{1}{2}$. We have

$$\mathscr{E}(a_1) = (t-1)\pi_{1.},$$

$$\operatorname{var} a_1 = \sum_{j=2}^{t} \pi_{1j}(1-\pi_{1j}) \leqslant \tfrac{1}{4}(t-1).$$

Since the variance of a_1 is a maximum under H_0, we may conclude that the above test will be conservative if H_0 is not true. In fact, as

$$\operatorname{var} a_1 \leqslant (t-1)\pi_{1.}(1-\pi_{1.})$$

we can also test the hypothesis that $\pi_{1.}$ has a specified value π_0 by referring a_1 to the binomial distribution with parameters $t-1$, π_0.

(ii) *Test of the equality of scores of two specified players*

We shall find the probability distribution $p(d)$ of $d = a_1 - a_2$, assuming H_0 to hold. From (3·3)

$$p(d) = \sum_{a_2} f(a_2+d, a_2) = \frac{1}{2^{2t-3}}\left[\Sigma\binom{t-2}{a_2+d-1}\binom{t-2}{a_2} + \Sigma\binom{t-2}{a_2+d}\binom{t-2}{a_2-1}\right],$$

where in view of the convention (3·4), the summations may be taken over all non-negative integers a_2. Thus

$$p(d) = \frac{1}{2^{2t-3}}\left[\binom{2t-4}{t-d-1} + \binom{2t-4}{t-d-3}\right], \quad d = -(t-1), \quad -(t-2)\ldots, \quad (t-1),$$

is a symmetrical distribution whose variance is $\frac{1}{2}t$. It can be shown that the characteristic function of d is

$$\phi(u) = (\cos\tfrac{1}{2}u)^{2t-4}\cos u,$$

and hence that $d/\sqrt{(\frac{1}{2}t)}$ tends to a unit normal distribution. Numerical work indicates this tendency to be rapid.

At the 5% level of significance, the smallest significant values of $|d|$ are as follows:

t	$\lvert d\rvert$
$\leqslant 4$	no significant values of $\lvert d\rvert$ possible
5–6	4
7–10	5
> 10	smallest integer $> 1{\cdot}96\sqrt{(\frac{1}{2}t)} + 0{\cdot}5$.

While this approach tests equality of the two scores under H_0 it will also provide a conservative test of the null hypothesis

$$H_0'\colon \quad \pi_{1.} = \pi_{2.} \quad \text{against the alternative } \pi_{1.} \neq \pi_{2.};$$

for, in the general case, we find

$$\operatorname{var} d = 4\pi_{12}(1-\pi_{12}) + \sum_{j=3}^{t}[\pi_{1j}(1-\pi_{1j}) + \pi_{2j}(1-\pi_{2j})],$$

which is a maximum under H_0.

(iii) *Test of the top score*

The test developed in this section is essentially different from those in (i) and (ii) in that it is not based on the scores of *specified* players, but on the score x_1 of the winning player (or players). We obtain, on H_0, the probability of large values of x_1 from which upper percentage points may be constructed. If in a particular tournament the appropriate upper percentage point is attained or exceeded by one or more players, H_0 is rejected and, as in the analogous problem of testing for 'outliers', these players may be regarded as 'superior'.

It is easy to see that for $t \leqslant 8$ the probability distribution of x_1 could be built up from Table 1. We shall employ a different approach which is useful also for larger values of t. Since x_1 is the largest of the scores $a_1, a_2, \ldots, a_t$, the method of inclusion and exclusion gives (see, for example, David, 1956)

$$\Pr(x_1 \geqslant t-r) = \sum_{s=1}^{t} (-)^{s-1} \binom{t}{s} \Pr(a_1 \geqslant t-r, a_2 \geqslant t-r, \ldots, a_s \geqslant t-r), \tag{3·6}$$

where r is any integer from 1 to t. For small values of r, the later terms in the sum will be zero, since only a few players can have a large score. We illustrate the method on the evaluation of the typical term

$$\Pr(a_1 \geqslant t-3, a_2 \geqslant t-3, a_3 \geqslant t-3, a_4 \geqslant t-3).$$

Consider the sub-tournament played by C_1, C_2, C_3 and C_4. From Table 1 we see that the possible outcomes are

$$[3210], \quad [31^3], \quad [2^30], \quad [2^21^2]$$

with respective frequencies

$$24, \quad 8, \quad 8, \quad 24.$$

It is, however, clear that only $[31^3]$ and $[2^21^2]$ can result in final scores all greater than $t-4$. The possible outcomes (ignoring the order of C_1 to C_4), and the frequencies with which they can arise from $[31^3]$ and $[2^21^2]$, are as follows:

				$[31^3]$	$[2^21^2]$
$t-1$,	$t-3$,	$t-3$,	$t-3$	1	0
$t-2$,	$t-2$,	$t-3$,	$t-3$	0	1
$t-2$,	$t-3$,	$t-3$,	$t-3$	$t-4$	$2(t-4)$
$t-3$,	$t-3$,	$t-3$,	$t-3$	$\frac{1}{2}(t-4)(t-5)$	$(t-4)^2$

This gives a *total* frequency of $28t^2-172t+272$, which can be converted into the required probability on division by $2^{2(2t-5)}$, the factor of equation (3·5). In this way we obtain from (3·6)

$$\Pr(x_1 = t-1) = \frac{t}{2^{t-1}},$$

$$\Pr(x_1 \geqslant t-2) = \frac{t^2}{2^{t-1}} - \frac{\binom{t}{2}(t-1)}{2^{2t-4}} + \frac{\binom{t}{3}}{2^{3t-7}},$$

$$\Pr(x_1 \geqslant t-3) = \frac{t}{2^t}(t^2-t+2) - \frac{\binom{t}{2}}{2^{2t-3}}(t^3-4t^2+7t-4)$$
$$+ \frac{\binom{t}{3}}{2^{3t-6}}(5t^3-33t^2+78t-64) - \frac{\binom{t}{4}}{2^{4t-12}}(7t^2-43t+68) + \frac{3\binom{t}{5}}{2^{5t-18}}.$$

These three probabilities are as follows for $t \leqslant 20$:

t	Pr $(x_1 \geqslant t-1)$	Pr $(x_1 \geqslant t-2)$	Pr $(x_1 \geqslant t-3)$
3	0·7500	1·0000	1·0000
4	0·5000	1·0000	1·0000
5	0·3125	0·9766	1·0000
6	0·1875	0·8418	1·0000
7	0·1094	0·6447	0·9987
8	0·0625	0·4526	0·9637
9	0·0352	0·2989	0·8536
10	0·0195	0·1891	0·6868
11	0·0107	0·1161	0·5093
12	0·0059	0·0696	0·3548
13	0·0032	0·0410	0·2360
14	0·0017	0·0238	0·1517
15	0·0009	0·0137	0·0950
16	0·0005	0·0078	0·0584
17	0·0003	0·0044	0·0353
18	0·0001	0·0025	0·0211
19	0·0001	0·0014	0·0125
20	0·0000	0·0008	0·0073

We see from this table that an unbroken sequence of wins is not significant, at the 5 % level, until $t = 9$ in the case of a one-sided test and $t = 10$ for a two-sided test. For a one-sided test a score of $t-2$ is significantly high for $t \geqslant 13$, and a score of $t-3$ for $t \geqslant 17$.

4. Discussion

Some properties of Knock-out and Round Robin tournaments have been presented. We now make a few remarks about their respective merits and drawbacks.

The attraction of Knock-out tournaments is, at least in part, that a large number of players can be handled quickly. With $t = 2^p$ competitors a decision is reached after p rounds and $t-1$ games as against $t-1$ rounds and $\frac{1}{2}t(t-1)$ games required for a Round Robin tournament if no play-offs are necessary. A natural objection to Knock-out tournaments is that they are too unreliable. This is often true in the simple cases discussed in § 2, but in view of the relative brevity of such tournaments it may well be practicable to repeat them. The simplest way of doing this is to start with a fresh random draw for each repetition and to declare as winner the player with the largest total number of wins, possibly after a play-off. Alternatively, the initial draws of each repetition could be arranged to satisfy certain symmetry requirements. We do not recommend either of these approaches if the aim of the tournament is to rank all players in order of merit. For this purpose a simple Round Robin tournament will be preferable to $\frac{1}{2}t$ repetitions of a Knock-out tournament, although both require the same (minimum) number of games. However, the situation will be very different when we are primarily concerned with picking the best player, since the repeated Knock-out tournament will tend to include a larger number of encounters between the stronger players. Some work along these lines is in progress. There are also possibilities, for both types of

tournaments, of some sequential procedure whereby the number of contestants is progressively reduced.

In this paper attention has been concentrated on tournaments without repetition of any kind. When comparing a number of treatments in a paired comparison experiment, we will often require some repetition, possibly by different judges. If repeated Round Robin tournaments are impracticably large, the incomplete block type designs put forward by Kendall (1955) and Bose (1956) are helpful. However, to answer the type of question considered in § 3, generalizations of tests there given as well as new tests are required. It is hoped to deal with some of these issues in a future publication.

I should like to thank Dr R. A. Bradley for some helpful discussions and Dr I. R. Savage for a valuable comment. My thanks are also due to Miss Betty Laby and Miss G. E. Jacobs of the University of Melbourne and to Mr A. Romano of the Virginia Polytechnic Institute for assistance with the computations.

REFERENCES

BOSE, R. C. (1956). Paired comparison designs for testing concordance between judges. *Biometrika*, **43**, 113–21.

BRADLEY, R. A. (1954). Rank analysis of incomplete block designs. II. Additional tables for the method of paired comparisons. *Biometrika*, **41**, 502–37.

BRADLEY, R. A. & TERRY, M. E. (1952). Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika*, **39**, 324–45.

DAVID, H. A. (1956). On the application to statistics of an elementary theorem in probability. *Biometrika*, **43**, 85–91.

DURBIN, J. (1951). Incomplete blocks in ranking experiments. *Brit. J. Psych.* (Statistics Section), **4**, 85–90.

KENDALL, M. G. (1955). Further contributions to the theory of paired comparisons. *Biometrics*, **11**, 43–62.

KENDALL, M. G. & BABINGTON SMITH, B. (1940). On the method of paired comparisons. *Biometrika*, **31**, 324–45.

MOSTELLER, F. (1951). Remarks on the method of paired comparisons. I. The least squares solution assuming equal standard deviations and equal correlations. *Psychometrika*, **16**, 3–9.

TEICHROEW, D. (1956). Tables of expected values of order statistics for samples of size twenty and less from the normal distribution. *Ann. Math. Statist.* **27**, 410–26.

THURSTONE, L. L. (1927). Psychophysical analysis. *Amer. J. Psychol.* **38**, 368–89.

ESTIMATION OF THE NORMAL POPULATION PARAMETERS GIVEN A SINGLY CENSORED SAMPLE

By J. G. SAW

University College London

1. Introduction

The problem of estimating the parameters of a normal population given an ordered sample $x_1 < x_2 \ldots < x_n$ censored above x_r has been treated by several authors.

Gupta (1952) has investigated the maximum likelihood solutions and the most efficient linear unbiassed estimators for μ and σ, the location and scale parameters. The simultaneous equations of the maximum likelihood solution are not immediately soluble. In his paper Gupta provides a table which, if enlarged to give finer argument intervals, would make the solution comparatively rapid. No information, however, seems to be available as to the extent to which the resulting estimates of μ and σ are biased; this is likely to be of importance in small samples. The estimators proposed below, while based on the same statistics as those involved in the maximum likelihood equations, namely x_r, and the first and second moments of the observations about x_r, combine them in such a way that the estimates of μ and σ^2 are unbiassed, however small the sample.

The method of most efficient linear unbiased estimators, whilst more easily handled computationally, suffers from two defects. In order to obtain the system of weights, it is necessary to invert a matrix of order equal to the number (r) of observations available. These weights differ for every set of values taken by r and n, thus requiring extensive tabling and arduous preliminary computations. Further, and this is a more serious difficulty, the weights are functions of $\mathscr{E}(y_i)$, $\mathscr{E}(y_i^2)$ and $\mathscr{E}(y_i y_j)$, where $y_1 < y_2 \ldots < y_n$ is an ordered sample from the standard normal population. At present these are only tabled for sample sizes $n \leqslant 20$, (Sarhan & Greenberg, 1956).

Gupta has overcome the labour of the matrix inversion by proposing a modified linear estimator, the weights being simple functions of $\mathscr{E}(y_i)$ only. These modified estimators are reasonably efficient but we are again limited in the use of them by the fact that, as yet, the values of $\mathscr{E}(y_i)$ are not tabulated with sufficient accuracy for $n > 20$.

2. Estimators symmetrical in $x_1, x_2, \ldots, x_{r-1}$

If we take

$$\mu^* = \epsilon \bar{x}_{r-1} + (1-\epsilon)\, x_r, \quad \text{with} \quad \bar{x}_{r-1} = \sum_{i=1}^{r-1} x_i/(r-1),$$

$$\eta^* = \alpha \sum_{i=1}^{r-1} (x_i - x_r)^2 + \beta \left[\sum_{i=1}^{r-1} (x_i - x_r) \right]^2$$

as estimators of μ and σ^2, respectively, the ϵ being chosen to make μ^* unbiased and α and β being chosen to make η^* unbiassed and of minimum variance, then providing these estimators are reasonably efficient, they will have immediate advantage owing to the small number of weights which have to be established. In fact it is found that the large sample efficiency of μ^* varies between 0·948 when $r/(n+1) = 0{\cdot}20$ and 0·993 when $r/(n+1) = 0{\cdot}80$.

Further, μ^* has a smaller variance than the corresponding modified estimator due to Gupta. The asymptotic efficiency of η^* is unity for all values of $r/(n+1)$.

Apart from the fact that they occur in the maximum likelihood equations, the reason for considering such statistics as $\bar{x}_{r-1}$, $\sum_1^{r-1}(x_i-x_r)^2$ and $\left[\sum_1^{r-1}(x_i-x_r)\right]^2$ is that the values $x_1,\ldots,x_{r-1}$ enter symmetrically. Thus we may regard $x_1,\ldots,x_{r-1}$ as an ordered sample from the population conditional on x_r, with density function

$$P(x\mid x_r)=\frac{1}{\sigma\sqrt{(2\pi)}}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}\Big/\int_{-\infty}^{(x_r-\mu)/\sigma}\frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}u^2}\,du\quad(-\infty<x<x_r).$$

However, for the purposes of this work, since μ^* and η^* are symmetrical in $x_1,\ldots,x_{r-1}$, we need no longer regard them as ordered. In effect therefore we no longer need the individual values of $\mathscr{E}(y_i)$, $\mathscr{E}(y_i^2)$ and $\mathscr{E}(y_iy_j)$ mentioned in the introduction.

In a previous paper by the author (1958) it was shown how it was possible to obtain moments and product moments of $\bar{x}_{r-1}$ and x_r in terms of a linear sum of integrals of the type

$$\psi(p_r,n:a,b)=\frac{1}{B(r,n-r+1)}\int_{-\infty}^{\infty}Z(u)\,[F(u)]^{r-1}\,[1-F(u)]^{n-r}\left[\frac{Z(u)}{F(u)}\right]^a u^b\,du,$$

where $$Z(u)=\frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}u^2},\quad F(u)=\int_{-\infty}^{u}Z(t)\,dt\quad\text{and}\quad p_r=r/(n+1),$$

and also how it was possible to expand these integrals as a power series in $1/(n+2)$; tables were given for $a+b\leqslant 4$, $p_r=0{\cdot}50(0{\cdot}05)\,0{\cdot}80$. We may treat moments and product moments of $\Sigma(x_i-x_r)^2$ and $[\Sigma(x_i-x_r)]^2$ in a similar manner. Ultimately, therefore, ϵ, α and β may be expressed as power series in $1/(n+2)$ though they are given at the end of the paper as power series in $1/(n+1)$, since this form is thought to converge more rapidly. In addition, it is possible using Sarhan & Greenberg's tables to give exact values for ϵ and var (μ^*) for $1<r<n\leqslant 20$.

3. Estimation of μ

Writing

$$\mu'_{10}=\mathscr{E}[(\bar{x}_{r-1}-\mu)\mid\sigma],\quad \mu'_{01}=\mathscr{E}[(x_r-\mu)\mid\sigma],\quad \mu_{ij}=\sigma^{-(i+j)}\mathscr{E}[(\bar{x}_{r-1}-\mathscr{E}(\bar{x}_{r-1}))^i\,(x_r-\mathscr{E}(x_r))^j]$$

we see that if $\mu^*=\epsilon\bar{x}_{r-1}+(1-\epsilon)\,x_r$ is to be an unbiassed estimator for μ, then we must have

$$\epsilon=\mu'_{01}/(\mu'_{01}-\mu'_{10}).$$

With this value of ϵ

$$\text{var}\,(\mu^*)=\sigma^2(\mu'^2_{01}\mu_{20}-2\mu'_{01}\mu'_{10}\mu_{11}+\mu'^2_{10}\mu_{02})/(\mu'_{01}-\mu'_{10})^2.$$

Exact values of $\epsilon(r,n)$ and var (μ^*) are given in Table 1 for sample sizes $n\leqslant 20$.

For use with sample sizes greater than twenty, the expansion

$$\epsilon(r,n)=\sum_{i=0}^{5}E_i(p_r)\,(n+1)^{-i}+O(n+1)^{-6}\quad\text{with}\quad p_r=r/(n+1)$$

was obtained and will be found to give an excellent approximation to the true values at $n=19$ (correct to the eighth place of decimals in most cases). The weight ϵ is a function of both r and n; the complete form $\epsilon(r,n)$ has been used above in view of a relationship between $\epsilon(r,n)$ and $\epsilon(n-r+1,n)$ which is given below.

Similarly, we obtain an expansion for the variance of μ^* of the form

$$\operatorname{var}\left(\frac{\mu^*}{\sigma}\right)\Big|_{(r,n)} = \sum_{i=1}^{5} V_i(p_r)\,(n+1)^{-i} + O(n+1)^{-6}.$$

The quantities $E_i(p_r)$ and $V_i(p_r)$, which are functions of the ratio $p_r = r/(n+1)$ and i only, are given in Table 2 for $p_r = 0{\cdot}50\,(0{\cdot}05)\,0{\cdot}80$. Values to cover the interval $0{\cdot}20 \leqslant p_r < 0{\cdot}50$ may be obtained using the relations

$$\frac{r-1}{\epsilon(r,n)} + \frac{n-r}{\epsilon(n-r+1,n)} = n,$$

$$\operatorname{var}\left(\frac{\mu^*}{\sigma}\right)\Big|_{(r,n)} = \frac{1}{n} + \frac{\operatorname{var}(\mu^*/\sigma)\,|_{(n-r+1,n)} - (1/n)}{\{1 - n\epsilon(n+1-r,n)/(n-r)\}}.$$

4. Comparison of μ^* with other estimators

Gupta has tabled the large sample variance of $\hat{\mu}$, the maximum likelihood estimator of μ. Comparing these tables with the leading term $V_1(p_r)\,(n+1)^{-1}$, we obtain the asymptotic efficiency of μ^* for various values of p_r:

p_r	Asymptotic efficiency	p_r	Asymptotic efficiency
0·20	0·9485	0·55	0·9711
0·25	0·9479	0·60	0·9764
0·30	0·9491	0·65	0·9813
0·35	0·9519	0·70	0·9857
0·40	0·9559	0·75	0·9896
0·45	0·9606	0·80	0·9929
0·50	0·9658		

We may also compare the efficiency of μ^* and μ_G (Gupta's modified estimator), relative to μ_L, the most efficient linear unbiased estimator.
Thus for a sample of size ten we have

r	$\dfrac{\operatorname{var}\mu_L}{\operatorname{var}\mu^*}$	$\dfrac{\operatorname{var}\mu_L}{\operatorname{var}\mu_G}$
2	1·0000	1·0000
3	0·9819	0·9150
4	0·9756	0·8983
5	0·9771	0·9113
6	0·9825	0·9354
7	0·9890	0·9663
8	0·9946	0·9806
9	0·9985	0·9943

It should also be noticed that $\epsilon = 0$ whenever $p_r = r/(n+1) = \frac{1}{2}$. Thus if we are given the observations up to and including the median, we disregard all but the median value in forming an estimate of μ. Although this will at first sight seem a great waste of data, it will be observed from the first table in this section that the asymptotic efficiency of the median in this context is 96·58 %, whilst the ratio (variance μ_L)/(variance μ^*) takes the values:

0·9899 when $n = 5, r = 3$; 0·9843 when $n = 7, r = 4$;

and 0·9806 when $n = 9, r = 5$.

Thus when only $\frac{1}{2}(n+1)$ of the observations are available, very little is lost by neglecting all but the median observation as far as the estimate of μ is concerned. In contrast, the asymptotic efficiency of the median relative to the *complete* sample of n is 63·66 %.

5. Estimation of σ^2

Defining
$$H'_{ij} = \mathscr{E}\left\{\left[\sum_{l=1}^{r-1}(x_l - x_r)^2\right]^i \left[\sum_{l=1}^{r-1}(x_l - x_r)\right]^{2j}\right\} \sigma^{-2(i+j)},$$

we see that
$$\eta^* = \alpha \sum_{i=1}^{r-1}(x_i - x_r)^2 + \beta\left[\sum_{i=1}^{r-1}(x_i - x_r)\right]^2$$

is unbiassed for σ^2 and of minimum variance over all possible (α, β) if

$$\alpha = (H'_{10}H'_{02} - H'_{01}H'_{11})/(H'^2_{10}H'_{02} + H'^2_{01}H'_{20} - 2H'_{10}H'_{01}H'_{11}),$$
$$\beta = (H'_{01}H'_{20} - H'_{10}H'_{11})/(H'^2_{10}H'_{02} + H'^2_{01}H'_{20} - 2H'_{10}H'_{01}H'_{11}),$$

and that with these values of α and β

$$\operatorname{var}(\eta^*/\sigma^2) = \frac{H'_{20}H'_{02} - H'^2_{11}}{H'^2_{10}H'_{02} + H'^2_{01}H'_{20} - 2H'_{10}H'_{01}H'_{11}} - 1.$$

Proceeding as indicated in § 2, we find expansions of the form

$$\alpha = \sum_{i=1}^{4} A_i(p_r)(n+1)^{-i} + O(n+1)^{-5},$$

$$\beta = \sum_{i=2}^{5} B_i(p_r)(n+1)^{-i} + O(n+1)^{-6},$$

$$\operatorname{var}(\eta^*/\sigma^2) = \sum_{i=1}^{4} W_i(p_r)(n+1)^{-i} + O(n+1)^{-5}.$$

Values of $A_i(p_r)$, $B_i(p_r)$ and $W_i(p_r)$ are given in Table 3.

By comparing $W_1(p_r)(n+1)^{-1}$ with the large sample form of the variance of $\hat{\sigma}^2$, the maximum likelihood estimator of σ^2, it is found that η^* has asymptotic efficiency of 100 % for all values of the ratio $r/(n+1)$.

It has not been found possible to give simple relations connecting $\alpha(r,n)$, $\beta(r,n)$, $\operatorname{var}\eta^*|_{(r,n)}$ and $\alpha(n-r+1,n)$, $\beta(n-r+1,n)$, $\operatorname{var}\eta^*|_{(n-r+1,n)}$ without bringing in auxiliary tables. However, it is hoped that the existing range on p_r will meet most requirements. In any case, an estimate of variance will be most unreliable when more than half the observations have been censored.

6. Interpolation in Tables 2 and 3

Owing to the labour involved in producing the tables, it was not found possible to give a fine entry for p_r. However, the following method of interpolation gives satisfactory results.

Suppose that n and r are given and that $p_r = r/(n+1)$ is not a tabulated value. Let $U(0{\cdot}65)$ denote, for example, $\sum_{i=1}^{5} E_i(0{\cdot}65)\,(n+1)^{-i}$; then if $h = 20r/(n+1) - 13$, we may use as an approximation to $U(p_r)$ the expression

$$\frac{h(h^2-1)(h^2-4)(h^2-9)}{720}\left[\frac{U(0{\cdot}50)}{h+3} - \frac{6U(0{\cdot}55)}{h+2} + \frac{15U(0{\cdot}60)}{h+1} - \frac{20U(0{\cdot}65)}{h} + \frac{15U(0{\cdot}70)}{h-1} - \frac{6U(0{\cdot}75)}{h-2} + \frac{U(0{\cdot}80)}{h-3}\right].$$

This corresponds to Lagrange 7-point interpolation and is reasonably quick to apply.

As an example, using this method with $r = 7$, $n = 10$, we obtain

$\epsilon = 0{\cdot}3869\,5262$, whereas the true value from Table 1 is $0{\cdot}3869\,5268$;
$\mathrm{var}\,(\mu^*/\sigma) = 0{\cdot}1179\,6547$, whereas the true value from Table 1 is $0{\cdot}1179\,6680$;
$\alpha = 0{\cdot}1938$, $-\beta = 0{\cdot}01767$, $\mathrm{var}\,(\eta^*/\sigma^2) = 0{\cdot}3815$ — No check is available on these values but they are thought to be correct to the number of places of decimals given.

The formula was also used for extrapolation and applied to the power series for ϵ in the cases $n = 19$; $r = 17$, 18 and 19. Checking against the true values in Table 1, the errors were found to be 1.10^{-5}, 3.10^{-5} and 2.10^{-4}, respectively.

7. A worked example

The following example is taken from Gupta's paper.

The table gives the day on which the first seven of ten tested mice died after being inoculated with a uniform culture of human tuberculosis.

Day	41	44	46	54	55	58	60
Log day $= (x_i)$	1·613	1·644	1·663	1·732	1·740	1·763	1·778
Frequency	1	1	1	1	1	1	1

It is found that

$$\bar{x}_6 = 1{\cdot}6925, \quad x_7 = 1{\cdot}778, \quad \sum_{i=1}^{6}(x_i - x_7)^2 = 0{\cdot}062191, \quad \left[\sum_{i=1}^{6}(x_i - x_7)\right]^2 = 0{\cdot}263169.$$

Thus
$$\mu^* = (0{\cdot}3870)\,(1{\cdot}6925) + (0{\cdot}6130)\,(1{\cdot}778) = 1{\cdot}745,$$
$$\eta^* = (0{\cdot}1938)\,(0{\cdot}062191) - (0{\cdot}01767)\,(0{\cdot}263169) = 0{\cdot}00740.$$

An estimate of the variance of μ^* is $(0{\cdot}1180)\eta^* = 0{\cdot}873 \times 10^{-3}$.

An estimate of the variance of η^* is $(0{\cdot}3815)\,\eta^{*2} = 0{\cdot}209 \times 10^{-4}$.

In passing, it is perhaps worth mentioning that μ^* has been found to be distributed very close to normality. Thus for rough tests on the hypothesis $\mu = \mu_0$, it may be possible to refer $(\mu^* - \mu_0)/\sqrt{(\text{est. var}\,\mu^*)}$ to the standard normal distribution. It seems intuitively more likely, however, that the distribution of this ratio will be better represented by the Pearson

Type IV curve. Similarly, for a rough test of the hypothesis $\sigma = \sigma_0$, we could refer η^* to the Pearson Type III distribution with the appropriate mean and standard deviation. These two problems have yet to be fully investigated and will probably form the subject of a later paper.

I wish to express my gratitude to Prof. E. S. Pearson, whose suggestions are responsible for several improvements on the original paper.

REFERENCES

GUPTA, A. K. (1952). *Biometrika*, **39**, 260–73.
SARHAN, A. E. & GREENBERG, B. G. (1956). *Ann. Math. Statist.* **27**, 427–51.
SAW, J. G. (1958). *Biometrika*, **45**, 211–21. Corrigenda pp. 587–8.

Table 1. *Values of the weight ϵ and the variance of μ^*/σ for $1 < r < n \leqslant 20$*

		Weight, ϵ	var μ^*/σ
$n = 3$,	$r = 2$	Zero	0·4486 7110
$n = 4$,	$r = 2$	−0·4055 5159	0·5129 9280
	3	·3093 2088	·2882 4812
$n = 5$,	$r = 2$	−0·7411 0682	0·6111 2918
	3	Zero	·2868 3366
	4	·4724 9056	·2185 7232
$n = 6$,	$r = 2$	−1·0260 6712	0·7186 4005
	3	−0·2676 7985	·3041 5115
	4	·2226 9148	·2089 5708
	5	·5735 0999	·1774 4443
$n = 7$,	$r = 2$	−1·2733 1715	0·8263 6223
	3	−0·5023 8193	·3302 2028
	4	Zero	·2104 4686
	5	·3642 6439	·1674 8290
	6	0·6422 3186	0·1498 1238
$n = 8$,	$r = 2$	−1·4915 3218	0·9309 9884
	3	−0·7109 1496	·3604 2823
	4	− ·1997 0421	·2180 7919
	5	·1737 4521	·1646 3021
	6	0·4623 9456	0·1409 3561
	7	·6920 0548	·1298 1932
$n = 9$,	$r = 2$	−1·6867 5767	1·0312 9527
	3	−0·8983 5837	0·3924 5050
	4	− ·3802 8184	·2292 5388
	5	Zero	·1661 0128
	6	0·2960 5268	0·1368 8836
	7	·5344 6005	·1221 7529
	8	·7297 0989	·1146 2569
$n = 10$,	$r = 2$	−1·8633 5148	1·1269 0294
	3	−1·0685 0751	0·4250 4749
	4	−0·5448 8270	·2425 0203
	5	− ·1592 4349	·1703 1365
	6	0·1423 7402	0·1359 7144
	7	·3869 5268	·1179 6680
	8	·5896 3408	·1080 8017
	9	·7592 5335	·1026 6399
$n = 11$,	$r = 2$	−2·0245 4630	1·2178 6597
	3	−1·2238 3557	0·4573 7320
	4	−0·6960 0994	·2569 4929
	5	− ·3060 1488	·1763 0723
	6	Zero	0·1371 6243
	7	0·2492 5991	·1160 7211
	8	·4572 0905	·1040 6917
	9	·6332 2913	·0970 3950
	10	·7830 2152	·0929 9029
$n = 12$,	$r = 2$	−2·1728 1803	1·3044 1143
	3	−1·3678 0736	0·4859 3821
	4	−0·8356 5351	·2720 5514
	5	− ·4420 2130	·1834 7202
	6	−0·1324 0995	0·1398 0858
	7	·1205 7511	·1158 5477
	8	·3325 5203	·1018 4129
	9	·5131 4930	·0933 4070
	10	·6685 3886	·0881 2541
	11	0·8324 6899	0·0847 2881
$n = 13$,	$r = 2$	−2·3100 7581	1·3868 1251
	3	−1·5009 4676	0·5208 4280
	4	−0·9654 0881	·2874 7620
	5	− ·5686 7675	·1914 0681
	6	−0·2560 4852	0·1434 7564
	7	Zero	·1167 9900
	8	·2152 6220	·1009 0906
	9	·3993 2369	·0910 3555
	10	·5587 4621	·0847 5964
	11	0·6977 1536	0·0807 6006
	12	·8188 8582	·0782 8340
$n = 14$,	$r = 2$	−2·4378 4620	1·4653 7120
	3	−1·6250 7121	0·5513 4148
	4	−1·0865 6888	·3029 9021
	5	−0·6871 5159	·1998 4054
	6	−0·3719 4277	0·1478 6561
	7	− ·1133 1319	·1186 0682
	8	·1045 5472	·1009 3894
	9	·2915 1408	·0897 6990
	10	·4540 6085	·0825 0407
	11	0·5966 2366	0·0777 1198
	12	·7222 2478	·0745 6212
	13	·8327 4358	·0725 5809

Table 1 (*cont.*)

		Weight, ϵ	var μ^*/σ
$n = 15, r =$	2	−2·5573 5741	1·5403 7895
	3	−1·5910 5633	0·5162 8847
	4	−1·2001 9350	·3184 5172
	5	−0·7984 1909	·2085 8619
	6	−0·4809 6831	0·1527 6888
	7	− ·2201 2308	·1210 5385
	8	Zero	·1016 9465
	9	·1893 1561	·0892 9543
	10	·3543 9077	·0810 9450
	11	0·4997 5262	0·0755 6299
	12	·6285 8591	·0718 0372
	13	·7431 0067	·0692 6842
	14	·8446 4790	·0676 1792
$n = 16, r =$	2	−2·6696 1409	1·6121 0956
	3	−1·8506 3947	0·6097 5853
	4	−1·3071 6059	·3337 6521
	5	−0·9032 9419	·2175 1772
	6	−0·5838 7075	0·1580 3503
	7	− ·3210 9497	·1239 7744
	8	− ·0990 2945	·1030 4166
	9	·0922 8706	·0894 3219
	10	·2594 7023	·0803 4191
	11	0·4071 0746	0·0741 1147
	12	·5384 7038	·0697 8395
	13	·6559 1509	·0667 6886
	14	·7610 9259	·0646 9077
	15	·8549 8345	·0633 1093
$n = 17, r =$	2	−2·7754 4828	1·6808 1522
	3	−1·9538 0346	0·6376 7089
	4	−1·4082 0482	·3488 6837
	5	−1·0024 6439	·2265 2675
	6	−0·6812 8676	0·1635 5464
	7	− ·4168 0820	·1272 5750
	8	− ·1930 4540	·1047 4082
	9	Zero	·0900 4658
	10	·1689 7949	·0801 0670
	11	0·3185 2094	0·0732 1078
	12	·4519 4855	·0683 4595
	13	·5716 9627	·0648 8380
	14	·6795 4769	·0624 2044
	15	·7767 5745	·0606 9067
	16	0·8640 4026	0·0595 2219
$n = 18, r =$	2	−2·8755 5681	1·7467 2583
	3	−2·0514 7109	0·6647 4230
	4	−1·5039 4667	·3637 2129
	5	−1·0965 1402	·2355 6194
$n = 18, r =$	6	−0·7737 6245	0·1692 4581
	7	− ·5077 6661	·1308 0404
	8	− ·2825 0321	·1064 6866
	9	− ·0879 4308	·0910 3752
	10	·0825 9311	·0802 8344
	11	0·2337 6240	0·0727 5085
	12	·3689 2384	·0673 7399
	13	·4905 5815	·0634 8909
	14	·6005 1947	·0606 6704
	15	·7001 8374	·0586 2267
	16	0·7905 1758	0·0571 6368
	17	·8720 4112	·0561 6310
$n = 19, r =$	2	−2·9705 2895	1·8100 4988
	3	−2·1441 9951	0·6910 0073
	4	−1·5949 1464	·3780 2346
	5	−1·1859 4337	·2445 7027
	6	−0·8617 6868	0·1750 4878
	7	− ·5944 0913	·1345 4868
	8	− ·3678 0590	·1091 3104
	9	− ·1719 0350	·0923 2730
	10	Zero	·0807 9098
	11	0·1525 8385	0·0726 4730
	12	·2892 3182	·0667 8006
	13	·4124 5536	·0624 9205
	14	·5241 5083	·0593 3084
	15	·6257 5867	·0569 9385
	16	0·7183 5714	0·0552 7424
	17	·8026 9911	·0540 2945
	18	·8791 5996	·0531 6423
$n = 20, r =$	2	−3·0608 6752	1·8709 7592
	3	−2·2324 6540	0·7164 7801
	4	−1·6815 6241	·3925 8856
	5	−1·2711 8372	·2535 1721
	6	−0·9457 1364	0·1809 1765
	7	− ·6771 1925	·1384 3893
	8	− ·4493 0972	·1116 5405
	9	− ·2522 0650	·0938 5532
	10	− ·0790 8873	·0815 6600
	11	0·0747 4052	0·0728 3429
	12	·2126 8352	·0664 9576
	13	·3372 7331	·0618 2136
	14	·4504 3378	·0583 3664
	15	·5536 4389	·0557 2300
	16	0·6480 4129	0·0537 6121
	17	·7344 8217	·0522 9780
	18	·8135 5787	·0512 2505
	19	·8860 9203	·0504 6728

Table 2. *Coefficients of* $(n+1)^{-i}$ *in the series approximation to* (*a*) *the weight factor* ϵ *and* (*b*) *the variance, of the estimator of location,* μ^*

(*a*) $\epsilon = \sum_{i=0}^{5} E_i(p_r)\,(n+1)^{-i} + O(n+1)^{-6}$ with $p_r = r/(n+1)$

p_r	$E_0(p_r)$	$E_1(p_r)$	$E_2(p_r)$	$E_3(p_r)$	$E_4(p_r)$	$E_5(p_r)$
0·50	Zero	Zero	Zero	Zero	Zero	Zero
0·55	0·1486 5772	0·0783 132	0·006 281	− 0·041 07	− 0·0201	0·077
0·60	·2823 5913	·1375 787	·001 435	− ·077 85	− ·0248	·161
0·65	0·4034 0719	0·1818 639	− 0·012 337	− 0·113 38	− 0·0159	0·273
0·70	·5135 6217	·2138 610	− ·034 313	− ·150 44	·0100	·448
0·75	·6141 8208	·2352 568	− ·065 129	− ·192 53	·0646	·763
0·80	·7063 1072	·2468 760	− ·107 286	− ·245 66	·1784	1·435

(*b*) $\text{var}\left(\frac{\mu^*}{\sigma}\right) = \sum_{i=1}^{5} V_i(p_r)\,(n+1)^{-i} + O(n+1)^{-6}$

p_r	$V_1(p_r)$	$V_2(p_r)$	$V_3(p_r)$	$V_4(p_r)$	$V_5(p_r)$
0·50	1·5707 963	0·896 605	0·098 73	− 0·4265	− 0·132
0·55	1·4169 573	·716 520	·072 37	− ·1662	·275
0·60	1·3034 728	·633 049	·176 56	·2218	·970
0·65	1·2186 755	0·605 297	0·327 05	0·5913	1·481
0·70	1·1547 730	·610 085	·485 54	·9072	1·840
0·75	1·1064 195	·634 026	·638 59	1·1676	2·090
0·80	1·0698 709	·669 408	·782 04	1·4575	2·333

Table 3. *Coefficients of $(n+1)^{-i}$ in the series approximation to the weights α and β used in estimating σ^2 and for the variance of this estimator*

p_r	$A_1(p_r)$	$A_2(p_r)$	$A_3(p_r)$	$A_4(p_r)$
0·50	2·482906	5·61883	6·8137	16·886
0·55	2·213248	4·34260	4·3813	10·188
0·60	1·989519	3·40329	2·7795	6·215
0·65	1·800852	2·69428	1·7316	3·852
0·70	1·639499	2·14687	0·9989	2·226
0·75	1·499774	1·71526	0·4764	1·276
0·80	1·377385	1·36769	0·0451	0·515

p_r	$B_2(p_r)$	$B_3(p_r)$	$B_4(p_r)$	$B_5(p_r)$
0·50	−1·517094	−9·70572	−11·8257	−32·504
0·55	−1·603473	−7·85637	− 8·9535	−21·737
0·60	−1·604646	−6·42606	− 6·7706	−14·903
0·65	−1·560119	−5·30056	− 5·1900	−10·557
0·70	−1·491833	−4·40321	− 3·9500	− 7·335
0·75	−1·412189	−3·67797	− 2·9867	− 5·184
0·80	−1·328224	−3·08317	− 2·1429	− 3·438

p_r	$W_1(p_r)$	$W_2(p_r)$	$W_3(p_r)$	$W_4(p_r)$
0·50	4·965812	9·05846	16·1363	32·610
0·55	4·426496	7·11698	11·2447	20·516
0·60	3·979038	5·66638	7·9504	13·174
0·65	3·601703	4·55496	5·6512	8·494
0·70	3·278998	3·68410	4·0015	5·446
0·75	2·999548	3·00114	2·8366	3·585
0·80	2·754769	2·41742	1·8451	1·902

$$\alpha = \sum_{i=1}^{4} A_i(p_r)\,(n+1)^{-i} + O(n+1)^{-5}, \quad \beta = \sum_{i=2}^{5} B_i(p_r)\,(n+1)^{-i} + O(n+1)^{-6},$$

$$\text{var}\,(\eta^*/\sigma^2) = \sum_{i=1}^{4} W_i(p_r)\,(n+1)^{-i} + O(n+1)^{-5}, \quad \text{with} \quad p_r = r/(n+1).$$

ON HOTELLING'S GENERALIZATION OF T^2

By K. C. SREEDHARAN PILLAI† and PABLO SAMSON, Jr‡

The Statistical Center, University of the Philippines

1. Introduction

Let S_1 and S_2 denote the matrices consisting of the sum of squares and products of deviations from means for two independent samples of sizes n_1 and n_2, respectively, taken from p-variate normal populations. Then T_0^2 (Hotelling, 1951) may be defined by

$$T_0^2/(n_2-1) = \text{trace}\,(S_2^{-1}S_1).$$

It is a generalization of his T^2 statistic. Since the trace of a matrix equals the sum of its characteristic roots, $T_0^2/(n_2-1)$ is the sum of the characteristic roots of the matrix $S_2^{-1}S_1$. Define $U^{(s)}$ as the sum of the characteristic roots of $S_2^{-1}S_1$, where s is the number of non-zero roots. Thus $s = p$ if $n_1, n_2 > p$. It may be noted that $0 \leqslant U^{(s)} < \infty$.

We give here expressions for the moments of $U^{(s)}$ for $s = 2, 3, \ldots, 6$. Upper 5 and 1 % points of the distribution of $U^{(s)}$ have been computed for $s = 2, 3$ and 4 for various values of two supplementary parameters using the exact moment quotients of $U^{(s)}$. Comparisons of these values for $s = 2$ are made with those obtained from the exact distribution, and those from an approximation by Pillai (1954). For $s = 3$ and 4, comparisons are made only with the approximation. A numerical example is also given to illustrate the use of the criterion $U^{(s)}$.

2. The moments of $U^{(s)}$

By making use of a recurrence relation given by Pillai (1954, 1956) the first four central moments of $U^{(s)}$ for $s = 2, 3, 4, 5$ and 6 have been obtained. We shall write

$$m = \tfrac{1}{2}(n_1-s-2), \quad n = \tfrac{1}{2}(n_2-s-2). \tag{2·1}$$

The nature of the first three moments suggests the following general forms:

$$\mu_1^{(s)} = \frac{s\{m+\frac{1}{2}(s+1)\}}{n}, \tag{2·2}$$

$$\mu_2^{(s)} = \frac{s\{m+\frac{1}{2}(s+1)\}\{n+m+\frac{1}{2}(s+1)\}(2n+s)}{n^2(n-1)(2n+1)}, \tag{2·3}$$

$$\mu_3^{(s)} = \frac{2s\{m+\frac{1}{2}(s+1)\}\{n+m+\frac{1}{2}(s+1)\}(n+2m+s+1)(2n+2s)(2n+s)}{n^3(n-1)(n-2)(2n+2)(2n+1)}. \tag{2·4}$$

Although the fourth raw moment of $U^{(s)}$ has also been generalized (Pillai & Mijares, 1959), the central moment has not yet been obtained because of the heavy algebraic computations

† United Nations Senior Advisor in Mathematical Statistics and Visiting Professor of Statistics, Statistical Center, University of the Philippines.

‡ Lecturer in Statistics, Statistical Center, University of the Philippines; portions of this paper are based on the M.A.S. thesis of the latter author submitted to the Statistical Center, March 1958.

involved. The fourth central moments of $U^{(s)}$ for $s = 2, 3$ and 4 which are used for computation in the following sections are therefore given in their explicit forms:

$$\mu_4^{(2)} = \frac{A_{42}}{n^4(n-1)(n-2)(n-3)(2n+1)(2n-1)}, \tag{2·5}$$

where

$$A_{42} = 3(2m+3)(2n+2m+3)\{[4n^2+(2m+3)(2n+2m+3)(n+3)](n+1)(n-2) + 16n(n+2m+3)^2\},$$

$$\mu_4^{(3)} = \frac{A_{43}}{(n+1)n^4(n-1)(n-2)(n-3)(2n+1)(2n-1)}, \tag{2·6}$$

where

$$A_{43} = 9(m+2)(n+m+2)\{(n+1)(n-2)[-2n^2+(m+2)(n+m+2)\{12(n+4)^2 - 88(n+4)+125\}]+(n+2m+4)^2[8(n+4)^3-38(n+4)^2-24(n+4)+134]\},$$

$$\mu_4^{(4)} = \frac{A_{44}}{(n+1)n^4(n-1)(n-2)(n-3)(2n+3)(2n+1)(2n-1)}, \tag{2·7}$$

where

$$A_{44} = 24(2m+5)(2n+2m+5)\{(n+1)(n-2)[2n^2+(2m+5)(2n+2m+5)\{(n+5)^3 - 12(n+5)^2+44(n+5)-49\}]+(n+2m+5)^2[2(n+5)^4 - 17(n+5)^3+20(n+5)^2+111(n+5)-224]\}.$$

Hence,

$$\beta_1^{(s)} = \frac{4(n+2m+s+1)^2(2n+2s)^2(n-1)(2n+1)}{s\{m+\frac{1}{2}(s+1)\}\{n+m+\frac{1}{2}(s+1)\}(2n+s)(n-2)^2(2n+2)^2}, \tag{2·8}$$

$$\beta_2^{(2)} = \frac{(n-1)(2n+1)A_{42}}{(2m+3)^2(2n+2m+3)^2(n+1)^2(n-2)(n-3)(2n-1)}, \tag{2·9}$$

$$\beta_2^{(3)} = \frac{(n-1)(2n+1)A_{43}}{9(m+2)^2(n+m+2)^2(2n+3)^2(2n-1)(n+1)(n-2)(n-3)} \tag{2·10}$$

and

$$\beta_2^{(4)} = \frac{(n-1)(2n+1)A_{44}}{4(2m+5)^2(2n+2m+5)^2(n+2)^2(2n+3)(2n-1)(n+1)(n-2)(n-3)}. \tag{2·11}$$

It may be observed that the moments of $U^{(s)}$ can be obtained in a simple manner from those of $V^{(s)}$, another criterion which is defined as the sum of $\theta_i = \lambda_i/(1+\lambda_i)$. In the numerator of any moment of $V^{(s)}$, if one maintains the coefficient of the highest power of n positive and put alternatively positive and negative signs for subsequent powers of n, then substitute for n the value $m+n+s+1$, one obtains the numerator of the corresponding moment of $U^{(s)}$. Similarly, the denominator of any moment of $U^{(s)}$ can be derived from the corresponding moment of $V^{(s)}$ by putting positive coefficients for terms involving n and negative coefficients for other terms in each linear factor and making the substitution of $m+n+s+1$ for n.

3. THE UPPER PERCENTAGE POINTS OF THE DISTRIBUTION OF $U^{(2)}$

The moment quotients, $\beta_1^{(2)}$ and $\beta_2^{(2)}$, have been computed using (2·8) and (2·9) for parameters $m = 1(0·5)\,5(5)\,30$; 50 and $n = 15(5)\,50$; 60, 80, 100 (except for a few values of n and smaller values of m). The upper 5 and 1 % points are then evaluated from 'Tables of the percentage points of Pearson curves for given β_1 and β_2 expressed in standardized measure' (Pearson &

Hartley, 1954, Table 42) and are shown in Tables 4 and 5 printed at the end of this paper.

Hotelling (1931, 1951) has obtained the exact C.D.F. of $U^{(2)}$ in terms of incomplete β-functions and it is given below:

$$P = I_w(2m+2, 2n+3) - \frac{(2n+2m+4)!\left(\frac{1-w}{1+w}\right)^{n+1}}{2(2m+1)!\,(2n+2)!} B_{w^2}(m+1, n+2), \tag{3·1}$$

where $w = U^{(2)}/(2+U^{(2)})$.

Pillai (1954, 1956) has suggested an approximation to the distribution of $U^{(s)}$ in the form

$$p(U^{(s)}) = \frac{[U^{(s)}]^{s(m+\frac{1}{2}(s+1))-1}}{B\{s[m+\frac{1}{2}(s+1)], sn+1\}\, s^{s(m+\frac{1}{2}(s+1))}\left(1+\frac{U^{(s)}}{s}\right)^{s(n+m+\frac{1}{2}(s+1))+1}} \tag{3·2}$$

and the approximate C.D.F. for $s = 2$ can be shown to be

$$P = I_w(2m+3, 2n+1). \tag{3·3}$$

A comparison of percentage points from the moment quotients with those from the exact and approximate C.D.F.'s can be judged from Table 1. A study of these tables confirms Pillai's observation (1954) that the error in using his approximate distribution is quite negligible for larger values of n; for example, for the upper 5% points, the approximate distribution could be used when n is 30 or more; for the upper 1% points, when n is a little larger, say around 40. The percentage points from the moment quotients differ from those of the exact C.D.F. only in the third significant digits for large values of n, and appear relatively closer to the exact percentage points than the approximate percentage points. These observations justify the use of percentage points from the moment quotients or the use of the percentage points from the approximate C.D.F. for larger values of n.

Table 1. *Comparison of approximations to the upper percentage points of $U^{(2)}$*

m	n	5% points			1% points		
		From exact C.D.F.	From approx. C.D.F.	From β_1, β_2	From exact C.D.F.	From approx. C.D.F.	From β_1, β_2
0	5	1·43	1·39	—	2·17	2·05	—
0	10	0·671	0·664	—	0·979	0·931	—
4	10	1·927	1·886	—	2·516	2·415	—
0	15	0·443	0·435	—	0·624	0·602	—
4	15	1·236	1·219	—	1·563	1·526	—
0	20	0·328	0·323	—	0·455	0·443	—
3	20	0·7690	0·7600	0·7683	0·9698	0·9499	0·9702
3·5	20	0·839	0·829	0·842	1·050	1·029	1·059
4	20	0·908	0·898	0·907	1·130	1·108	1·130
5	20	1·046	1·034	1·045	1·288	1·263	1·288
0	30	0·2155	0·2124	—	0·2960	0·2904	—
3	30	0·5020	0·4983	0·5022	0·6235	0·6165	0·6236

4. UPPER PERCENTAGE POINTS OF THE DISTRIBUTION OF $U^{(s)}$ FOR CASES $s = 3$ AND 4

The moment quotients, β_1 and β_2 have been computed for various values of the parameters m and n using (2·8), (2·10) and (2·11) for $s = 3$ and 4. The percentage points were then again evaluated from Table 42 of Pearson & Hartley (1954) and are shown in Tables 6–9.

Making use of Pillai's approximate C.D.F. given by expression (3·2), we have for $s = 3$,

$$p = I_w(3m+6, 3n+1), \quad \text{where} \quad w = U^{(3)}/(3+U^{(3)}) \tag{4·1}$$

and for $s = 4$

$$p = I_w(4m+10, 4n+1), \quad \text{where} \quad w = U^{(4)}/(4+U^{(4)}). \tag{4·2}$$

For comparison with the upper 5 and 1 % points from the moment quotients, percentage points were also evaluated for certain values of m and n from the approximate C.D.F. given in (4·1) and (4·2). The results are tabulated together with the corresponding percentage points from β_1 and β_2 in Tables 2 and 3.

Table 2. *Comparison of approximations to the upper percentage points of* $U^{(3)}$

n	5 % points				1 % points			
	$m = 0$		$m = 5$		$m = 0$		$m = 5$	
	From β_1, β_2	From approx. C.D.F.	From β_1, β_2	From approx. C.D.F.	From β_1, β_2	From approx. C.D.F.	From β_1, β_2	From approx. C.D.F.
15	—	—	2·120	2·076	—	—	2·560	2·462
20	—	—	1·557	1·533	—	—	1·852	1·798
25	—	—	1·229	1·213	—	—	1·449	1·415
30	0·3621	0·3575	1·015	1·005	0·4635	0·4524	1·190	1·168
35	0·3097	0·3055	0·8638	0·8571	0·3939	0·3867	1·009	0·9943
40	·2692	·2671	·7520	·7483	·3412	·3376	0·8759	·8672
45	·2387	·2373	·6660	·6636	·3028	·2995	·7736	·7680
50	·2144	·2133	·5977	·5957	·2714	·2686	·6927	·6886
60	0·1778	0·1774	0·4959	0·4936	0·2250	0·2227	0·5728	0·5693
80	·1331	·1327	·3688	·3687	·1674	·1662	·4238	·4240
100	·1062	·1060	·2949	·2947	·1333	·1328	·3384	·3382

A study of these tables shows that the percentage points from the two methods are close for larger values of n and hence the approximate distributions appear usable as follows:

	5 %	1 %
$U^{(3)}$	$n \geqslant 40$	$n \geqslant 50$
$U^{(4)}$	$n \geqslant 50$	$n \geqslant 60$

Table 3. *Comparison of approximations to the upper percentage points of* $U^{(4)}$

n	5% points				1% points			
	$m = 0$		$m = 5$		$m = 0$		$m = 5$	
	From β_1, β_2	From approx. C.D.F.	From β_1, β_2	From approx. C.D.F.	From β_1, β_2	From approx. C.D.F.	From β_1, β_2	From approx. C.D.F.
15	1·122	1·084	2·877	2·812	1·413	1·334	3·395	3·246
20	0·8272	0·808	2·113	2·085	1·028	0·9841	2·458	2·387
25	·6548	·6415	1·668	1·611	0·8069	·7800	1·925	1·881
30	·5423	·5342	1·378	1·373	·6644	·6478	1·581	1·550
35	0·4629	0·4572	1·174	1·172	0·5646	0·5530	1·341	1·329
40	·4033	·3994	1·023	1·014	·4911	·4819	1·165	1·156
45	·3574	·3544	0·9058	0·9035	·4343	·4264	1·029	1·021
50	·3209	·3184	·8130	·8095	·3893	·3821	0·9218	0·9134
60	0·2666	0·2651	0·6745	0·6738	0·3224	0·3180	0·7627	0·7598
80	·1990	·1984	·5030	·5039	·2399	·2375	·5670	·5674
100	·1588	·1583	·4010	·4018	·1909	·1895	·4512	·4518

5. An example

The criterion $U^{(4)}$ is useful for testing different kinds of hypotheses in multivariate analysis, one being the comparison of two independent estimates of the same dispersion matrix. A second problem in which this criterion can be used is in testing the equality of p-dimensional mean vectors of l p-variate normal populations (which is mathematically identical with the general problem of multivariate analysis of variance of means). The values of m and n appropriate for this test are

$$m = \tfrac{1}{2}|l-p-1|-\tfrac{1}{2}, \quad n = \tfrac{1}{2}(N-l-p-1), \tag{5·1}$$

where N is the total of l sample sizes. The following data are taken from Rao (1952, p. 263) concerning measurements of 140 schoolboys of almost the same age belonging to six different schools in an Indian city. Three characters were measured, namely, (1) head length, (2) height and (3) weight. The problem is to test for the significance of differences between schools in the mean characters. Let S^*, S be the matrices consisting of the sums of squares and products 'between' and 'within' schools for the three characters. Then

$$S^* = \begin{array}{c} \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{pmatrix} 752{\cdot}0 & 214{\cdot}2 & 521{\cdot}3 \\ . & 151{\cdot}3 & 401{\cdot}2 \\ . & . & 1612{\cdot}7 \end{pmatrix} \end{array} \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} \tag{5·2}$$

$$S = \begin{array}{c} \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{pmatrix} 12{,}809{\cdot}3 & 1003{\cdot}7 & 2671{\cdot}2 \\ . & 1499{\cdot}6 & 4123{\cdot}6 \\ . & . & 21{,}009{\cdot}6 \end{pmatrix} \end{array} \begin{matrix} 1 \\ 2. \\ 3 \end{matrix} \tag{5·3}$$

The value of $U^{(3)} = 0{\cdot}2016$ is obtained as the trace of the matrix $S^{-1}S^*$, where S^{-1} denotes the reciprocal of the S matrix. Referring to Tables 6 and 7, with $m = 0{\cdot}5$ and $n = 65$, we find that this value of $U^{(3)}$ is just significant at the 5% level, which falls at about 0·195, and is clearly not significant at the 1% level, 0·242. The corresponding percentage points from Pillai's approximate c.d.f. are 0·194 and 0·239, respectively. This agrees with the findings of Rao, who examined the data using the Λ criterion of Pearson & Wilks (1933). Foster (1957), however, finds that the *largest root* is significant only at the 15% level.

The fact that in this particular example different test functions based on the same roots give different results need not be regarded as surprising. Far too little is known at present about alternatives to the null hypothesis to which these different tests are most sensitive.

The authors wish to acknowledge the facilities offered by the Statistical Center, University of the Philippines, during the course of preparation of this paper.

REFERENCES

Foster, F. G. (1957). *Biometrika*, **44**, 441.

Hotelling, H. (1931). *Ann. Math. Statist.* **2**, 360.

Hotelling, H. (1944). Air testing of sample bombsights. Prepared for the Applied Mathematics Panel, National Defense Research Committee, by the Statistical Research Group, Columbia University, *SRG Report No.* 328, *AMP Report No.* 5, 1*R*.

Hotelling, H. (1951). A generalized T test and measure of multivariate dispersion. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, **2**, 23.

Pearson, E. S. & Hartley, H. O. (1954). *Biometrika Tables for Statisticians*, **1**. Cambridge University Press for the Biometrika Trustees.

Pearson, E. S. & Wilks, S. S. (1933). *Biometrika*, **25**, 353.

Pillai, K. C. S. (1954). On some distribution problems in multivariate analysis. *Mimeographed Series No.* 88. Institute of Statistics, University of North Carolina.

Pillai, K. C. S. (1956). *Ann. Math. Statist.* **27**, 1106.

Pillai, K. C. S. & Mijares, T. A. (1959). *Ann. Math. Statist.* **30**.

Rao, C. R. (1952). *Advanced Statistical Methods in Biometric Research*. New York: John Wiley and Sons, Inc.

Table 4. *Upper 5 % points of the sum of two non-null roots, $U^{(2)}$*

m \ n	15	20	25	30	35	40	45	50	60	80	100
1·0	—	—	—	—	—	—	—	—	—	0·116	0·092
1·5	—	—	—	—	0·310	0·270	0·239	0·215	0·179	·133	·106
2·0	—	—	0·496	0·410	·350	·305	·270	·242	·202	·150	·120
2·5	—	—	·551	·456	·389	·339	·300	·269	·224	·167	·133
3·0	—	0·768	·607	·502	·428	·373	·330	·296	·246	·183	·146
3·5	—	0·842	0·662	0·547	0·466	0·406	0·360	0·323	0·268	0·200	0·159
4·0	—	·907	·717	·592	·505	·439	·389	·349	·290	·216	·172
4·5	—	·976	·771	·637	·542	·472	·418	·375	·311	·232	·185
5·0	—	1·045	·824	·681	·580	·505	·447	·401	·332	·248	·197
10	—	1·718	1·352	1·115	0·947	0·823	0·728	0·652	0·541	0·402	0·366
15	3·263	2·381	1·870	1·539	1·306	1·134	1·003	·899	·744	·553	·440
20	4·171	3·039	2·383	1·958	1·661	1·442	1·275	1·142	·945	·701	·558
25	5·077	3·692	2·894	2·374	2·014	1·749	1·545	1·383	1·143	·848	·674
30	5·981	4·345	3·403	2·792	2·366	2·053	1·813	1·622	1·341	·994	·790
50	9·590	6·947	5·430	4·449	3·767	3·266	2·881	2·577	2·127	1·573	1·248

Table 5. *Upper 1 % points of the sum of two non-null roots, $U^{(2)}$*

m \ n	15	20	25	30	35	40	45	50	60	80	100
1·0	—	—	—	—	—	—	—	—	—	0·148	0·118
1·5	—	—	—	—	0·408	0·344	0·304	0·273	0·226	·168	·134
2·0	—	—	0·632	0·520	·442	·384	·339	·303	·251	·187	·149
2·5	—	—	·697	·572	·485	·422	·372	·334	·276	·205	·163
3·0	—	0·970	·760	·624	·529	·458	·406	·364	·301	·224	·178
3·5	—	1·059	0·822	0·675	0·572	0·497	0·439	0·393	0·325	0·242	0·192
4·0	—	1·130	·884	·725	·616	·534	·472	·422	·349	·259	·206
4·5	—	1·210	·945	·776	·658	·571	·504	·451	·373	·277	·220
5·0	—	1·288	1·005	·825	·700	·607	·536	·480	·396	·294	·234
10	—	2·057	1·602	1·310	1·108	0·960	0·846	0·757	0·625	0·462	0·366
15	3·934	2·812	2·183	1·782	1·505	1·302	1·147	1·024	·845	·624	·495
20	4·974	3·551	2·758	2·243	1·895	1·640	1·443	1·288	1·061	·783	·621
25	6·043	4·302	3·330	2·712	2·286	1·974	1·736	1·549	1·275	·940	·745
30	7·092	5·043	3·900	3·173	2·673	2·306	2·028	1·807	1·488	1·097	·868
50	11·28	7·996	6·167	5·007	4·202	3·626	3·185	2·837	2·328	1·711	1·353

Table 6. *Upper 5 % points of the sum of three non-null roots,* $U^{(3)}$

m \ n	15	20	25	30	35	40	45	50	60	80	100
0·0	—	—	—	0·362	0·309	0·269	0·239	0·214	0·178	0·133	0·106
0·5	—	0·658	0·520	·431	·368	·321	·284	·255	·212	·158	·126
1·0	—	·761	·603	·499	·425	·371	·329	·295	·245	·183	·146
1·5	1·172	·864	·683	·566	·482	·420	·372	·334	·278	·207	·165
2·0	1·310	·964	·763	·632	·538	·469	·416	·373	·310	·231	·184
2·5	1·446	1·065	0·843	0·696	0·594	0·517	0·458	0·411	0·341	0·254	0·203
3·0	1·581	1·164	·921	·761	·649	·565	·500	·449	·372	·278	·222
3·5	1·717	1·263	·998	·825	·703	·612	·542	·486	·403	·293	·240
4·0	1·852	1·362	1·076	·889	·757	·659	·583	·524	·434	·324	·258
4·5	1·986	1·459	1·153	·952	·810	·706	·625	·561	·465	·347	·277
5·0	2·120	1·557	1·229	1·015	·864	·752	·666	·598	·496	·369	·295
10	3·438	2·517	1·982	1·634	1·391	1·210	1·071	0·961	0·796	0·593	0·472
15	4·742	3·494	2·725	2·245	1·933	1·661	1·469	1·316	1·090	·811	·646
20	6·039	4·407	3·478	2·852	2·424	2·107	1·862	1·668	1·381	1·026	·817
25	6·933	5·080	4·003	3·303	2·811	2·445	2·163	1·940	1·607	1·196	·953
30	8·168	5·981	4·711	3·886	3·306	2·875	2·543	2·278	1·888	1·406	1·120
50	10·10	7·327	5·734	4·708	3·990	3·460	3·053	2·731	2·255	1·671	1·326

N.B. For $m = -0{\cdot}5$, $n = 100$, the 5 % point of $U^{(3)}$ is 0·085.

Table 7. *Upper 1 % points of the sum of three non-null roots,* $U^{(3)}$

m \ n	15	20	25	30	35	40	45	50	60	80	100
0·0	—	—	—	0·464	0·394	0·341	0·303	0·271	0·225	0·167	0·110
0·5	—	0·837	0·658	·541	·459	·400	·353	·317	·262	·195	·156
1·0	—	·956	·750	·617	·524	·456	·403	·361	·299	·223	·177
1·5	1·475	1·071	·840	·691	·587	·511	·451	·404	·335	·249	·198
2·0	1·634	1·186	·930	·765	·649	·564	·499	·447	·370	·275	·219
2·5	1·791	1·298	1·018	0·837	0·711	0·617	0·546	0·489	0·405	0·301	0·240
3·0	1·947	1·410	1·106	·909	·771	·670	·592	·531	·439	·326	·260
3·5	2·102	1·521	1·193	·980	·831	·722	·638	·572	·473	·339	·279
4·0	2·252	1·632	1·278	1·050	·891	·774	·684	·612	·506	·376	·299
4·5	2·408	1·742	1·364	1·120	·950	·825	·729	·653	·540	·401	·319
5·0	2·560	1·852	1·449	1·190	1·009	·876	·774	·693	·573	·424	·338
10	4·060	2·925	2·283	1·871	1·583	1·372	1·211	1·084	0·895	0·664	0·528
15	5·541	4·026	3·101	2·536	2·180	1·858	1·639	1·465	1·210	·896	·712
20	7·009	5·030	3·938	3·196	2·701	2·338	2·061	1·842	1·520	1·124	·893
25	7·838	5·658	4·419	3·618	3·066	2·658	2·345	2·098	1·732	1·285	1·021
30	9·212	6·639	5·184	4·244	3·593	3·114	2·747	2·455	2·029	1·505	1·196
50	11·69	8·318	6·432	5·237	4·406	3·803	3·344	2·982	2·451	1·805	1·428

Table 8. *Upper 5 % points of the sum of four non-null roots,* $U^{(4)}$

m \ n	10	15	20	25	30	35	40	45	50	60	80	100
−0·5	—	—	0·712	0·547	0·453	0·387	0·337	0·299	0·268	0·223	0·167	0·133
0·0	—	1·122	·827	·655	·542	·463	·403	·357	·321	·267	·199	·159
0·5	—	1·302	·961	·761	·630	·537	·468	·415	·372	·309	·231	·184
1·0	—	1·482	1·092	·865	·716	·610	·531	·471	·426	·351	·262	·209
1·5	—	1·660	1·223	·968	·800	·682	·594	·526	·472	·392	·293	·234
2·0	—	1·837	1·352	1·120	·884	·753	·656	·581	·522	·433	·324	·258
2·5	—	2·012	1·480	1·171	0·967	0·824	0·718	0·636	0·571	0·474	0·354	0·282
3·0	—	2·186	1·608	1·271	1·050	·895	·780	·691	·620	·515	·384	·306
3·5	—	2·359	1·735	1·371	1·132	·963	·841	·745	·669	·555	·414	·330
4·0	—	2·532	1·862	1·470	1·214	1·035	·902	·799	·717	·595	·444	·354
4·5	—	2·710	1·988	1·569	1·296	1·104	·962	·852	·765	·636	·473	·378
5·0	—	2·877	2·113	1·668	1·378	1·174	1·023	·906	·813	·674	·503	·401
10	7·17	4·578	3·353	2·644	2·183	1·855	1·618	1·430	1·284	1·064	0·793	0·632
15	9·84	6·266	4·581	3·610	2·978	2·533	2·203	1·949	1·748	1·448	1·078	·859
20	12·15	7·404	5·803	4·570	3·768	3·204	2·784	2·463	2·207	1·828	1·361	1·083
25	15·15	9·622	7·021	5·526	4·553	3·869	3·363	2·974	2·665	2·207	1·641	1·306
30	17·80	11·30	8·238	6·481	5·337	4·533	4·050	3·464	3·121	2·583	1·920	1·528
50	28·40	17·98	13·09	10·29	8·464	7·184	6·238	5·505	4·938	4·082	3·031	2·410

Table 9. *Upper 1 % points of the sum of four non-null roots,* $U^{(4)}$

m \ n	10	15	20	25	30	35	40	45	50	60	80	100
−0·5	—	—	0·908	0·687	0·565	0·480	0·418	0·370	0·331	0·274	0·204	0·163
0·0	—	1·413	1·028	·807	·664	·565	·491	·434	·389	·322	·240	·191
0·5	—	1·620	1·177	·925	·761	·647	·562	·497	·446	·369	·275	·219
1·0	—	1·824	1·325	1·040	·856	·727	·632	·559	·501	·415	·309	·246
1·5	—	2·025	1·470	1·154	·949	·807	·701	·620	·556	·460	·342	·272
2·0	—	2·224	1·615	1·322	1·042	·885	·769	·680	·609	·504	·375	·299
2·5	—	2·421	1·758	1·378	1·134	0·963	0·836	0·739	0·662	0·548	0·408	0·325
3·0	—	2·618	1·899	1·489	1·225	1·039	·902	·798	·715	·592	·440	·350
3·5	—	2·813	2·039	1·599	1·314	1·116	·969	·856	·767	·635	·472	·376
4·0	—	3·008	2·180	1·708	1·402	1·191	1·034	·914	·819	·678	·504	·401
4·5	—	3·202	2·319	1·817	1·493	1·266	1·100	·972	·870	·722	·535	·426
5·0	—	3·395	2·458	1·925	1·581	1·341	1·165	1·029	·922	·763	·567	·451
10	8·57	5·305	3·827	2·991	2·452	2·075	1·803	1·591	1·425	1·179	0·875	0·696
15	11·66	7·189	5·177	4·038	3·309	2·802	2·430	2·144	1·919	1·585	1·175	·934
20	14·74	9·065	6·520	5·079	4·159	3·521	3·050	2·690	2·406	1·986	1·472	1·169
25	17·81	10·94	7·857	6·115	5·004	4·233	3·666	3·233	2·890	2·385	1·767	1·403
30	20·88	12·80	9·19	7·15	5·85	4·94	4·40	3·74	3·37	2·78	2·06	1·63
50	33·13	20·26	14·51	11·27	9·20	7·77	6·72	5·91	5·29	4·36	3·22	2·55

TABLES FOR WALD TESTS FOR THE MEAN OF A NORMAL DISTRIBUTION

By ELIZABETH D. BARRACLOUGH and E. S. PAGE

Durham University Computing Laboratory

1. Introduction

If x is a normally distributed random variable with unknown mean θ and known standard deviation σ, the Wald sequential test for the hypothesis $\theta = \theta_0$ against the alternative $\theta = \theta_1$ ($\theta_1 > \theta_0$) is specified as follows (Wald, 1947, pp. 117 *et seq.*):

Continue taking observations x_i as long as

$$a < \frac{\theta_1 - \theta_0}{\sigma^2} \sum_{i=1}^{n} x_i + \frac{n}{2\sigma^2}(\theta_0^2 - \theta_1^2) < b; \tag{1}$$

if the left-hand inequality first fails to be satisfied, stop sampling and accept $\theta = \theta_0$, similarly accept $\theta = \theta_1$ if the right-hand inequality fails.

In practice the test is usually carried out either by plotting the cumulative sum Σx_i on a chart on which the parallel straight line boundaries are drawn and stopping sampling when one of the boundaries has been crossed or by comparing the cumulative sum after each observation with acceptance and rejection numbers calculated from (1).

If
$$y_i = \{2x_i - (\theta_0 + \theta_1)\}/2\sigma, \tag{2}$$
the inequalities (1) may be written
$$-Z < \sum_{i=1}^{n} y_i < h - Z, \tag{3}$$
where
$$Z = -\frac{a\sigma}{\theta_1 - \theta_0}, \quad h = \frac{(b-a)\,\sigma}{\theta_1 - \theta_0}. \tag{4}$$

The variate y_i is normally distributed with unit standard deviation and mean
$$\mu = \{2\theta - (\theta_0 + \theta_1)\}/2\sigma, \tag{5}$$
where $E(x) = \theta$.

Let $P(Z \mid \mu)$ be the probability of accepting the hypothesis that $\theta = \theta_0$ when the true mean has a particular value θ, and μ is given by equation (5). The argument μ will be omitted where no ambiguity is possible. The derivation of the equation for $P(Z)$ given by Anscombe (1948) considers expectations conditional upon the first observation (Page, 1954*a*) and in this case yields
$$P(Z) = \Phi(-\mu - Z) + \int_0^h P(x)\,(2\pi)^{-\frac{1}{2}} \exp -\tfrac{1}{2}(x - Z - \mu)^2\,dx, \tag{6}$$
where
$$\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}t^2}\,dt.$$

If $\mu = \mp \frac{1}{2}(\theta_1 - \theta_0)/\sigma$, $P(Z \mid \mu)$ gives the probabilities of accepting $\theta = \theta_0$ when $E(x) = \theta_0$ and $E(x) = \theta_1$, respectively. The charts and Tables 1 *a–d*, of the Appendix enable h and Z to be selected to give specified values for these two probabilities. The constants of the test, a and b, can then be calculated from equations (4). In Tables 1 *a–d*,

$$P_- = P(Z \mid -\tfrac{1}{2}(\theta_1 - \theta_0)/\sigma) = 1 - \alpha,$$
$$P_+ = P(Z \mid \tfrac{1}{2}(\theta_1 - \theta_0)/\sigma) = \beta.$$

If the required values of the operating characteristic are not among those tabulated, it may be sufficient for practical purposes to interpolate by eye in the charts provided. The values in the tables were those requested for a particular application, but the range covered is sufficiently wide to be of more general use. Further combinations of α and β can be obtained by using the relationships

$$P(h-Z \mid -\tfrac{1}{2}(\theta_1-\theta_0)/\sigma) = 1-P(Z \mid \tfrac{1}{2}(\theta_1-\theta_0)/\sigma),$$
$$P(h-Z \mid \tfrac{1}{2}(\theta_1-\theta_0)/\sigma) = 1-P(Z \mid -\tfrac{1}{2}(\theta_1-\theta_0)/\sigma).$$

Tables 2*a–d*, of the Appendix, computed from equation (9) below, give the average sample numbers, N_- (when $\theta = \theta_0$) and N_+ (when $\theta = \theta_1$) of each one of these tests.

2. The method of use of the charts and tables

(*a*) *To select a test with given risks*

If a test of the mean of a normal distribution with standard deviation σ is required which has probability $1-\alpha$ of accepting the hypothesis $\theta = \theta_0$ when $\theta = \theta_0$, and probability β when $\theta = \theta_1$ (so that the risks of the first and second kinds are α, β, respectively) it may be found as follows.

Calculate μ_0, μ_1 where $$\mu_0 = -\mu_1 = (\theta_0-\theta_1)/2\sigma. \tag{7}$$

Charts are given for $\mu = \pm 0{\cdot}25$, $\pm 0{\cdot}5$, $\pm 0{\cdot}75$, $\pm 1{\cdot}0$, and there are curves for $P(Z \mid \mu > 0) = 0{\cdot}05$, $0{\cdot}10$, $0{\cdot}20$, $0{\cdot}30$, $0{\cdot}40$, $0{\cdot}50$, $0{\cdot}60$, $0{\cdot}70$ and $P(Z \mid \mu < 0) = 0{\cdot}95$, $0{\cdot}99$, $0{\cdot}995$, $0{\cdot}999$.

Example 1

To select a test with $P(Z \mid \theta_0 = 3{\cdot}0) = 0{\cdot}99$, $P(Z \mid \theta_1 = 3{\cdot}75) = 0{\cdot}05$ when $\sigma = 0{\cdot}5$.

We have $$\mu_0 = -\mu_1 = (3{\cdot}0-3{\cdot}75)/2(0{\cdot}5) = -0{\cdot}75.$$

From Table 1*c* for $\mu = \pm 0{\cdot}75$, $h = 3{\cdot}88$; $Z = 1{\cdot}42$.

Hence $a = -2{\cdot}13$, $b = 3{\cdot}69$ from (4). The required test is therefore to take observations as long as $$-0{\cdot}71+3{\cdot}375n < \Sigma x_i < 1{\cdot}23+3{\cdot}375n.$$

(*b*) *Operating characteristics*

Several points on the operating characteristic (o.c.) may be obtained approximately by visual interpolation between the curves on the charts for the different values of μ.

Example 2

The approximate o.c. for the test selected in Example 1 is shown in Fig. 1 in which the circled points are those read from the charts.

(*c*) *Average sample number*

The average numbers of observations required for the tests given in Table 1 when the mean is that specified by either hypothesis are shown in Table 2.

Example 3

The average number of observations required by the test of Example 1 when $\theta = 3{\cdot}0$ is $N = 2{\cdot}87$ and when $\theta = 3{\cdot}75$ is $N = 4{\cdot}00$.

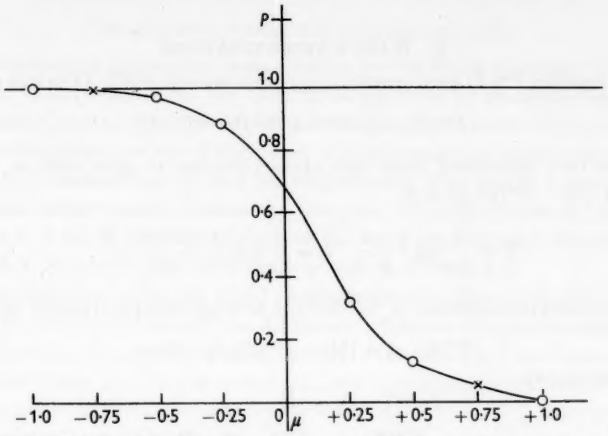


Fig. 1. Operating characteristic. × ×, defining points of the test. ○○, points read from the charts.

3. CALCULATION OF THE TABLES AND CHARTS

For a given pair of values of h and μ the solution $P(Z)$ to (6) was computed by an iterative method (Page, 1954*b*; Good, 1957). The known function, the normal distribution function, was calculated by a rational function approximation (Hastings, 1952) with an error at most $\frac{1}{2} \times 10^{-6}$. The integral containing the unknown solution was replaced by a Gaussian quadrature formula using the smallest number of ordinates that would give a fairly accurate approximation:

$$P(Z) = \Phi(-\mu - Z) + \sum_{i=1}^{n} w_i P(x_i) K(x_i, Z), \tag{8}$$

where w_i, x_i are respectively the Gauss weights and abscissae, and $K(x, Z)$ is the kernel of (6). Equation (8) was used iteratively to obtain $P(Z)$, for $Z = x_i$, $i = 1, \ldots, n$, starting from an initial approximation (usually zero). The iteration continued until successive values of the $P(x_i)$ all differed by less than a small quantity. At this stage equation (8) was applied for $Z = y_j$, $j = 1, \ldots, n'$, where y_j are the abscissae of a Gauss quadrature formula with more ordinates, $n' > n$. The $P(y_j)$ so obtained formed the starting value for iterations with the more accurate formula. This process continued automatically until the $P(Z)$ values at a set of Gauss points were reproduced to a prescribed accuracy by successive quadrature formulae. Values of $P(Z)$ at other points were calculated and used by inverse interpolation to obtain estimated Z values for the given $P(Z)$. These estimates were refined by substitution in (8) and repeating the process.

Some of the pivotal values of the $P(Z)$ were checked by an independent solution of (8) by the finite difference methods of Fox & Goodwin (1953). Limitations of storage space prevented attainment of the same accuracy of evaluation of the integral, but in no case did the corresponding pivotal values differ by more than one unit in the fourth decimal place from those derived by the first method.

The average sample number $N(Z)$ satisfies the equation

$$N(Z) = 1 + \int_0^h N(x)\,(2\pi)^{-\frac{1}{2}} \exp -\tfrac{1}{2}(x - Z - \mu)^2\,dx, \tag{9}$$

which was solved by the first method for the tests quoted in Table 1.

4. Wald's approximations

Wald's approximation $P^*(Z)$ for the operating characteristic $P(Z)$ of the test may be written

$$P^*(Z) = (e^{-2\mu h} - e^{-2\mu Z})/(e^{-2\mu h} - 1); \tag{10}$$

accordingly the test calculated from this approximation to give risks α, β, i.e. so that $P^*(Z \mid -\mu) = 1-\alpha$, $P^*(Z \mid \mu) = \beta$

has
$$Z = \frac{1}{2\mu}\log\frac{1-\alpha}{\beta}, \quad h = \frac{1}{2\mu}\log\frac{(1-\alpha)(1-\beta)}{\alpha\beta}. \tag{11}$$

Wald also gives an approximation $N^*(Z)$ for the average sample number of the test.

$$N^*(Z \mid \mu) = [h\{1 - P^*(Z|\mu)\} - Z]/\mu. \tag{12}$$

Thus for the test above

$$N^*(Z \mid -\mu) = (Z - h\alpha)/\mu, \tag{13}$$

$$N^*(Z \mid \mu) = \{h(1-\beta) - Z\}/\mu. \tag{14}$$

A comparison of $P^*(Z \mid \mu)$ and $P(Z \mid \mu)$ for one value of h, two values of Z and a range of μ was made by Page (1954*a*). The present table allows a more extensive comparison. Text-table 1 shows the Wald approximations α^*, β^* to the risks calculated from (10) for some tests with risks α, β at means $\pm\mu$. For $\mu = 0{\cdot}25$ the approximate risks of the second kind, β, are about one-third greater than their correct values, while for $\mu = 1$ they are more than three times greater. Further, for $\mu = 1$, no approximations are possible when $\alpha = 0{\cdot}001$, $\beta = 0{\cdot}3$, for example, since $Z < 0$ and (10) produces a result greater than unity. A similar behaviour is shown for risks of the first kind. It is clear that for any work requiring an accurate assessment of the risks for $\mu \geqslant 0{\cdot}25$ the Wald approximation cannot be accepted. Further, some of the Wald approximations exceed the true O.C. for $\mu < 0$ if Z is small; these are not likely to occur for the most frequently used schemes with small α and β.

Text-table 1

μ	h	Z	True O.C.	Wald O.C.	True A.S.N.	Wald A.S.N. with Wald O.C.	Wald A.S.N. with true O.C.
−0·25	11·77	0·37	0·999	0·9994	4·32	1·4	1·4
− ·25	14·90	5·40	·995	·992	23·84	21·1	21·3
− ·25	10·61	5·31	·95	·934	21·46	18·4	19·1
+0·25	11·77	0·37	0·60	0·831	16·99	6·5	17·4
+ ·25	14·90	5·40	·05	·067	37·43	34·0	35·0
+ ·25	10·61	5·31	·05	·066	21·46	18·4	19·1
−1·0	3·31	0·48	0·999	0·9979	1·49	0·5	0·5
−1·0	1·72	0·86	·95	·848	1·63	0·6	0·8
+1·0	3·31	0·48	0·10	0·382	3·24	1·6	2·5
+1·0	1·72	0·86	·05	·152	1·63	0·6	0·8

Comparison of exact and approximate risks

The above comparison shows that tests selected by the Wald formulae to have specified risks will in fact usually have risks less than those specified by amounts depending largely upon the population mean. It would be useful to make an extensive comparison of the exact average sample numbers (A.S.N.) of the exact and approximate tests with specified risks, i.e. those with h, Z, chosen from (6) and (8), respectively. This would, however, entail more computation than is warranted. Instead we compare the approximate A.S.N. for the exact test with values h, Z taken from the tables and the approximate A.S.N. for the approximate test with values h, Z given by (9) with the true values of the A.S.N.

The approximate A.S.N. using the Wald approximation to the O.C. is always smaller than the true value for the exact test and the difference is frequently considerable. If, however, the true O.C. is used in (12) a better approximation to the true A.S.N. is obtained, but it is still too small in most cases and in particular no such approximation is useful if $Z < 0$. For $\mu = 0{\cdot}25$, this approximation seems to yield an error of less than 10 %, but as μ increases the approximation deteriorates.

Improved approximations to the O.C. and A.S.N. have been derived by modifying the integral equations (6) and (9) (Page, 1954*a*) and more recently by a similar method by Kemp (1958). In the examples given, Kemp's approximations to $P(0)$, $N(0)$ are usually nearer to the true values than Page's* but the latter are closer approximations for a range of Z near the middle of $(0, h)$. No further comparisons of these approximations have been made, nor of their quality over the wider range of schemes now available.

We wish to thank Dr O. L. Davies and the Pharmaceuticals Division of Imperial Chemical Industries for their financial support for the preparation of the graphs.

REFERENCES

ANSCOMBE, F. J. (1948). Cambridge lecture notes.
FOX, L. & GOODWIN, E. T. (1953). *Phil. Trans.* A, **245**, 501–34.
GOOD, I. J. (1957). *Math. Tab., Wash.*, **11**, 82–3.
HASTINGS, L. J. (1952). *Rational Approximations in Numerical Analysis.* Rand Corporation, Santa Monica. (Subsequently published by Princeton University Press, 1955.)
KEMP, K. W. (1958). *J. Roy. Statist. Soc.* B, **20**, 379–86.
PAGE, E. S. (1954*a*). *J. Roy. Statist. Soc.* B, **16**, 136–9.
PAGE, E. S. (1954*b*). *Proc. Camb. Phil. Soc.* **50**, 414–25.
WALD, A. (1947). *Sequential Analysis.* New York: John Wiley and Sons.
WETHERILL, G. B. (1957). Ph.D. Thesis. University of London.

* Wetherill (1957) has found that the values of $N(0)$ for $k = 0{\cdot}25$, $0{\cdot}50$ in Table 4 of Page (1954*a*) have been miscalculated and that the corrections give better approximations to the true values.

APPENDIX

Table 1

P_- / P_+	0·999		0·995		0·99		0·95	
	h	*Z*	*h*	*Z*	*h*	*Z*	*h*	*Z*
	(*a*) $\mu = 0{\cdot}25$							
0·70	10·80	−0·03	7·57	−0·04	6·17	−0·05	2·81	−0·19
·60	11·77	0·37	8·55	0·36	7·15	0·35	3·82	0·24
·50	12·63	0·78	9·40	0·78	8·01	0·76	4·70	0·67
·40	13·45	1·25	10·23	1·24	8·84	1·23	5·53	1·15
0·30	14·22	1·83	11·12	1·82	9·72	1·81	6·42	1·73
·20	14·98	2·64	12·19	2·63	10·80	2·62	7·50	2·53
·10	—	—	13·78	4·01	12·42	4·00	9·12	3·92
·05	—	—	14·90	5·40	13·87	5·39	10·61	5·31
	(*b*) $\mu = 0{\cdot}5$							
0·70	4·42	−0·71	2·79	−0·72	2·08	−0·74	0·30	−0·95
·60	5·02	−0·39	3·40	−0·40	2·69	−0·42	0·97	−0·54
·50	5·54	−0·09	3·93	−0·10	3·22	−0·11	1·54	−0·19
·40	6·04	0·22	4·42	0·22	3·72	0·21	2·05	0·15
0·30	6·55	0·57	4·93	0·57	4·23	0·56	2·57	0·51
·20	7·13	1·02	5·51	1·02	4·81	1·01	3·16	0·97
·10	7·95	1·73	6·34	1·72	5·64	1·72	3·99	1·68
·05	8·69	2·42	7·08	2·41	6·38	2·41	4·73	2·37
	(*c*) $\mu = 0{\cdot}75$							
0·70	2·11	−1·12	1·01	−1·15	0·56	−1·18	—	—
·60	2·59	−0·82	1·50	−0·84	1·03	−0·86	—	—
·50	3·02	−0·55	1·93	−0·56	1·46	−0·57	0·27	−0·69
·40	3·42	−0·27	2·34	−0·28	1·87	−0·29	0·71	−0·37
0·30	3·83	0·03	2·75	0·03	2·28	0·02	1·14	−0·04
·20	4·28	0·40	3·20	0·40	2·74	0·39	1·62	0·35
·10	4·90	0·94	3·82	0·94	3·36	0·93	2·25	0·90
·05	5·42	1·42	4·34	1·42	3·88	1·42	2·78	1·39
	(*d*) $\mu = 1{\cdot}0$							
0·70	0·85	−1·45	—	—	—	—	—	—
·60	1·26	−1·16	0·47	−1·20	0·11	−1·24	—	—
·50	1·64	−0·90	0·83	−0·92	0·47	−0·94	—	—
·40	2·00	−0·63	1·19	−0·64	0·83	−0·66	—	—
0·30	2·36	−0·34	1·55	−0·35	1·20	−0·36	0·25	−0·44
·20	2·77	−0·00	1·97	−0·01	1·61	−0·01	0·69	−0·07
·10	3·31	0·48	2·51	0·48	2·16	0·47	1·27	0·44
·05	3·75	0·89	2·95	0·89	2·60	0·89	1·72	0·86

Table 2

P_+ \ P_-	0·999		0·995		0·99		0·95	
	N_-	N_+	N_-	N_+	N_-	N_+	N_-	N_+
	(*a*) $\mu = 0{\cdot}25$							
0·70	3·15	12·00	3·00	8·16	2·84	6·53	2·00	2·92
·60	4·32	16·99	4·15	11·86	3·97	9·67	2·97	4·72
·50	5·75	22·34	5·56	15·92	5·37	13·17	4·23	6·90
·40	7·53	28·00	7·32	20·31	7·11	17·00	5·84	9·41
0·30	9·82	33·72	9·61	25·05	9·38	21·18	7·97	12·27
·20	13·06	39·10	12·83	30·17	12·58	25·74	11·00	15·51
·10	—	—	18·34	35·69	18·06	30·82	16·22	19·27
·05	—	—	23·84	37·43	23·54	33·51	21·46	21·46
	(*b*) $\mu = 0{\cdot}5$							
0·70	1·29	3·35	1·25	2·39	1·21	1·99	1·04	1·11
·60	1·49	4·53	1·45	3·25	1·41	2·71	1·18	1·49
·50	1·77	5·83	1·73	4·22	1·68	3·53	1·41	1·97
·40	2·16	7·22	2·11	5·29	2·05	4·47	1·74	2·57
0·30	2·69	8·71	2·63	6·47	2·58	5·50	2·22	3·27
·20	3·48	10·32	3·42	7·74	3·35	6·64	2·96	4·08
·10	4·86	12·06	4·80	9·16	4·72	7·91	4·26	5·02
·05	6·25	13·00	6·18	9·95	6·10	8·63	5·58	5·58
	(*c*) $\mu = 0{\cdot}75$							
0·70	1·05	1·80	1·04	1·38	1·02	1·21	—	—
·60	1·10	2·28	1·08	1·71	1·07	1·47	—	—
·50	1·18	2·81	1·16	2·10	1·14	1·79	1·03	1·12
·40	1·29	3·40	1·27	2·55	1·25	2·18	1·12	1·35
0·30	1·47	4·05	1·44	3·05	1·42	2·62	1·27	1·63
·20	1·76	4·75	1·73	3·61	1·71	3·11	1·53	1·97
·10	2·33	5·52	2·30	4·23	2·27	3·68	2·06	2·39
·05	2·94	5·94	2·91	4·59	2·87	4·00	2·64	2·64
	(*d*) $\mu = 1{\cdot}0$							
0·70	1·01	1·29	—	—	—	—	—	—
·60	1·02	1·52	1·01	1·19	1·00	1·04	—	—
·50	1·04	1·79	1·03	1·38	1·02	1·21	—	—
·40	1·07	2·09	1·06	1·61	1·05	1·40	—	—
0·30	1·13	2·44	1·12	1·87	1·11	1·63	1·03	1·09
·20	1·24	2·82	1·23	2·17	1·21	1·89	1·12	1·27
·10	1·49	3·24	1·48	2·52	1·46	2·20	1·35	1·49
·05	1·80	3·48	1·78	2·71	1·76	2·38	1·63	1·63

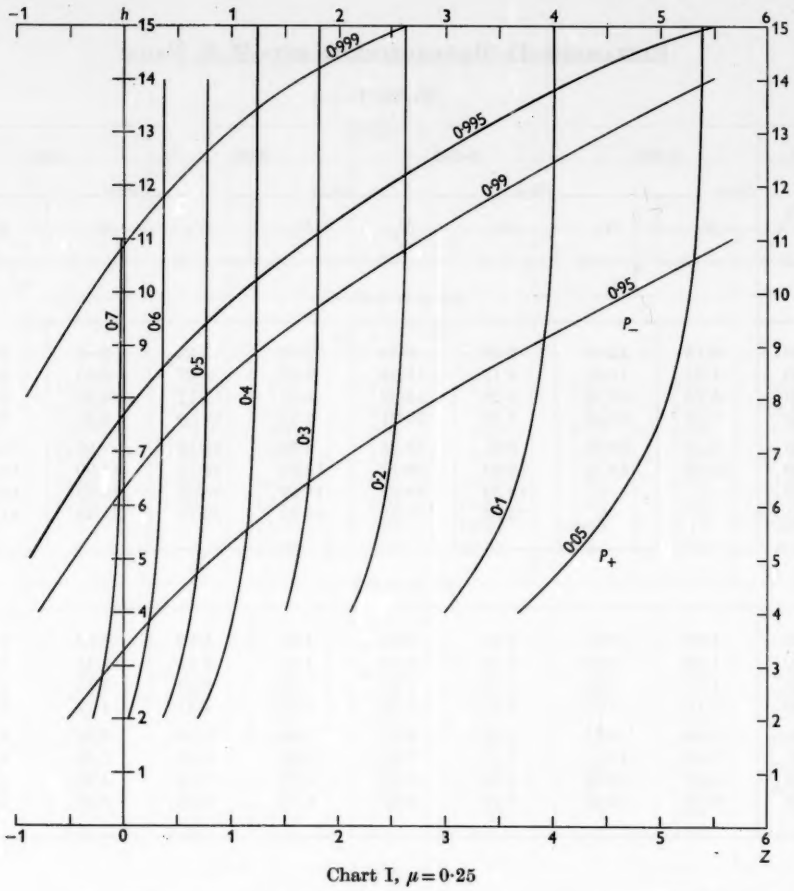


Chart I, $\mu = 0{\cdot}25$

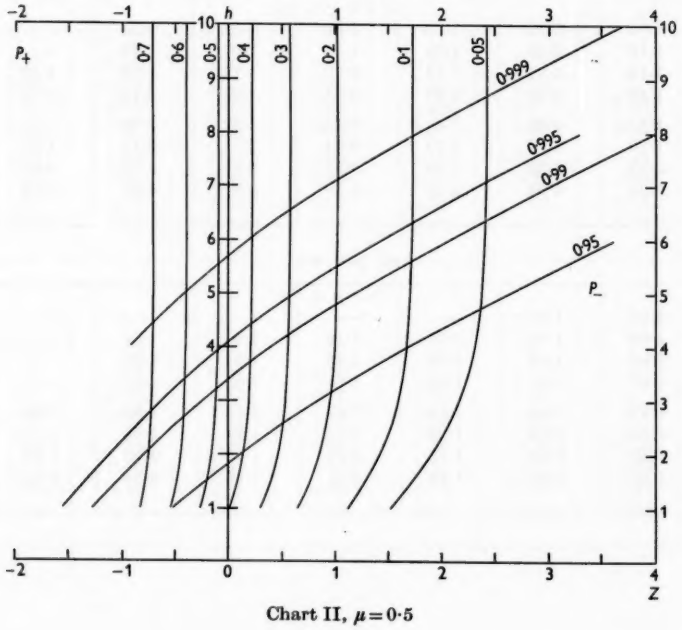


Chart II, $\mu = 0{\cdot}5$

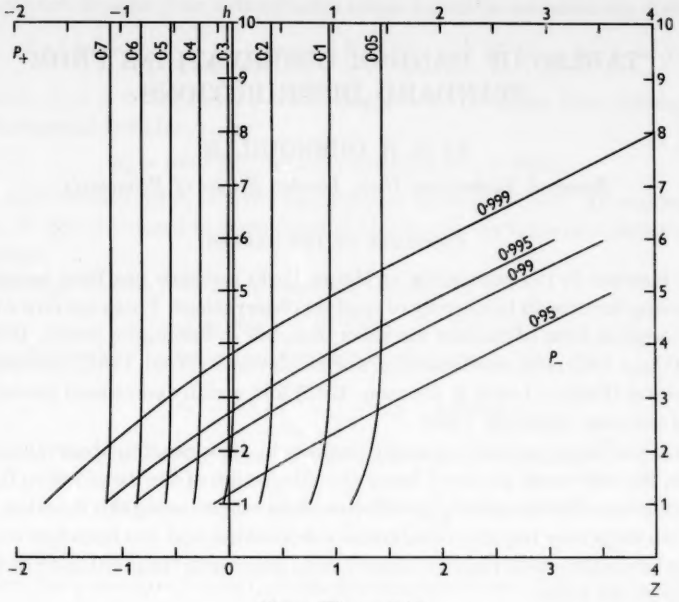


Chart III, $\mu = 0{\cdot}75$

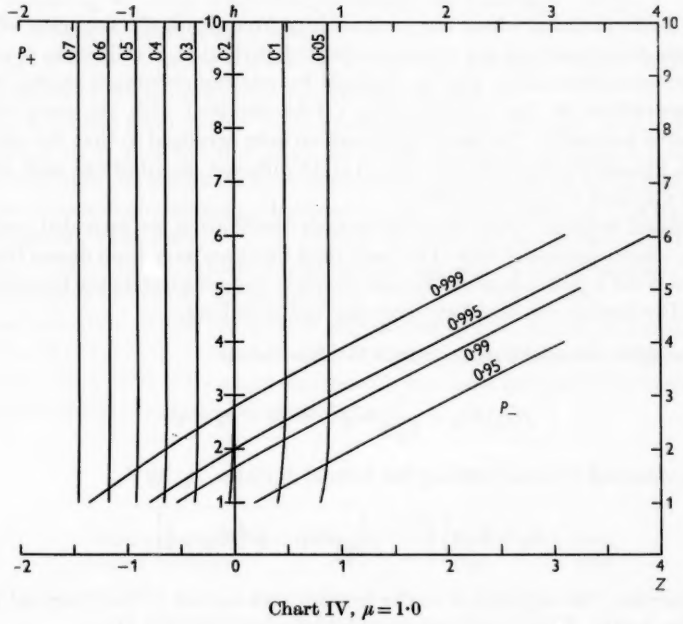


Chart IV, $\mu = 1{\cdot}0$

TABLES OF RANDOM OBSERVATIONS FROM STANDARD DISTRIBUTIONS

By M. H. QUENOUILLE

Research Techniques Unit, London School of Economics

Purposes of the tables

The recent increase in the popularity of Monte Carlo methods has been accompanied by a corresponding increase in the supply of random observations. These are now available not only in the original form of random numbers (Kendall & Babington Smith, 1939; Tippett, 1927; Rand Corp. 1955) but also as random normal deviates (Wold, 1954), correlated random normal deviates (Fieller, Lewis & Pearson, 1955) and serially correlated random numbers and normal deviates (Kendall, 1949).

From these published tables, it is easily possible to draw random observations from any distribution, the only steps involved being the calculation of the distribution function and the transformation of rectangularly distributed observations using this function. Nevertheless, these two steps may require considerable calculations, and it is therefore convenient to have readily available sets of random observations from some standard distributions. These are provided in this table.

Contents of the tables

For Monte Carlo methods where the purpose is to investigate the sampling effects of departures from the normal (or any other standard) distribution, it is likely in most instances that the greatest information will be obtained by relating estimated values for the non-normal observations to the corresponding values obtained with the *same* observations transformed to normality. The table has therefore been arranged so that the values in each line provide identical percentiles for each of eight different distributions each with a mean of zero and a second moment of unity.

One thousand random observations from each distribution are provided, and these are arranged in twenty groups of fifty. The basic data for these have been drawn from the first two pages of Wold's (1954) *Random Normal Deviates*, and are given in the first column, x_1, of each page. The further columns have been derived as follows.

(*a*). x_2, *random rectangular deviates from the distribution*

$$f(x_2)\,dx_2 = \frac{1}{2\sqrt{3}}\,dx_2, \quad -\sqrt{3} \leqslant x_2 \leqslant \sqrt{3}. \tag{1}$$

These were obtained by transforming the normal deviates, x_1, by

$$x_2 = 2\sqrt{3}\left[\int_{-\infty}^{x_1} \frac{1}{\sqrt{(2\pi)}} \exp -\tfrac{1}{2}x_1^2\,dx_1 - \tfrac{1}{2}\right].$$

In this calculation, the expression in the bracket was carried to four decimal places and multiplied by 3·4641. The answer was rounded off to two decimal places.

(*b*). x_3, *random deviates from a distribution whose logarithm was normally distributed*

$$f(x_3)\,dx_3 = \frac{1}{\sqrt{(2\pi)\,(x_3+a)}} \exp -\tfrac{1}{2}[\log_e(x_3+a)+b]^2\,dx_3, \quad (x_3 \geqslant -a) \tag{2}$$

where $a = (e-1)^{-\frac{1}{2}} = 0{\cdot}76287$ and $b = 0{\cdot}5 - \log_e a = 0{\cdot}771$. These were obtained by transforming the normal deviates, x_1, by

$$x_3 = (e^{x_1} - e^{\frac{1}{2}})/\sqrt{(e^2-e)} = 0{\cdot}46271\,e^{x_1} - 0{\cdot}76287.$$

This gives a lower limit for x_3 of $0{\cdot}763$ and a modal value of $-0{\cdot}593$. In making this transformation, e^{x_1} was calculated to three decimal places and the value of x_3 rounded off to two decimal places.

(*c*). x_4, *random deviates from the exponential distribution*

$$f(x_4)\,dx_4 = e^{-(x_4+1)}\,dx_4 \quad (x_4 \geqslant -1). \tag{3}$$

These were obtained by transforming the normal deviates, x_1, by

$$x_4 = -\log_e\left[\int_{x_1}^{\infty} \frac{1}{\sqrt{(2\pi)}} \exp -\tfrac{1}{2}x_1^2\,dx_1\right] - 1.$$

(*d*). x_5, x_6, x_7, *random deviates from an Edgeworth Type A expansion*

$$f(x)\,dx = \exp\left[-\frac{\kappa_3}{3!}D^3 + \frac{\kappa_4}{4!}D^4\right] \frac{1}{\sqrt{(2\pi)}} \exp -\tfrac{1}{2}x^2\,dx, \tag{4}$$

with x_5: $\kappa_3 = 0{\cdot}5$, $\kappa_4 = 0{\cdot}0$, with x_6: $\kappa_3 = 0{\cdot}0$, $\kappa_4 = 0{\cdot}5$ and with x_7: $\kappa_3 = 0{\cdot}5$, $\kappa_4 = 0{\cdot}5$.

These have been calculated using Cornish–Fisher expansions which give a sufficient degree of approximation within the central range of observations*. The expansions used, derived by substitution in formula (6·54) of Kendall and Stuart (1958), were

$$\begin{aligned}
124{,}416x_5 &= -9552 + 127{,}225x_1 + 7824x_1^2 - 40x_1^3 + 576x_1^4 - 252x_1^5,\\
1536x_6 &= 1411x_1 + 56x_1^3 - 3x_1^5,\\
124{,}416x_7 &= -12{,}144 + 122{,}878x_1 + 14{,}304x_1^2 - 1066x_1^3 - 720x_1^4 + 261x_1^5.
\end{aligned}$$

Tabulated values of these functions are given on page 182.

It is clear that these (or any other polynomial) expansions will break down for sufficiently extreme values of x_1 and the fact that tabulated values of x_5 reach a minimum near $x_1 = -2{\cdot}5$ indicates the unreliability of the series expansion for x_5 beyond this point. To preserve the monotonic relationship between x_1 and x_5, the values of x_5 corresponding to three more extreme values of x_1 ($-2{\cdot}79$, $-2{\cdot}76$, $-2{\cdot}69$) have been replaced by the minimum value for x_5, $-1{\cdot}86$. These replacements have been indicated with an asterisk. The original values, should they be needed, are respectively $-1{\cdot}81$, $-1{\cdot}82$, $-1{\cdot}84$.

A further limitation of the polynomial expansions is that values of cumulants of the distribution will not exactly equal the values of κ_3 and κ_4 inserted in the Edgeworth expansion. Thus the true values of the third and fourth cumulants (evaluated by taking expectations of powers of the polynomial expansions) are:

	x_5	x_6	x_7
Third cumulant	0·47	0·00	0·50
Fourth cumulant	0·01	0·45	0·56

* See note by E. S. Pearson on pp. 203–4 below.

(*e*). x_8, *random observations from the two-sided exponential distribution*

$$f(x_8)\,dx_8 = \frac{1}{2\sqrt{2}} \exp\left(-\frac{|x_8|}{\sqrt{2}}\right) dx_8 \quad (-\infty \leqslant x_8 \leqslant \infty). \tag{5}$$

These were obtained by transforming the normal deviates, x_1, by

$$x_8 = \frac{-\operatorname{sign} x_1}{\sqrt{2}} \left[\log_e \int_{|x_1|}^{\infty} \frac{1}{\sqrt{(2\pi)}} \exp -\tfrac{1}{2}x_1^2\, dx_1 + \log_e 2\right].$$

The values were calculated to four decimal places and rounded off to two decimal places.

All calculations were repeated (in many instances by the same computer) to provide a check. A further check was provided by punching each row of results on a Hollerith card and ordering the set of these according to values of x_1. All other variables should then be arranged in order.

This check revealed (apart from a number of errors in punching the cards) four or five omissions of minus signs. No gross error was found.

Properties of the distributions

The theoretical means and variances were, of course, zero and unity. The overall observed values showed close agreement with these theoretical values:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Mean	0·00	0·00	−0·02	−0·01	0·00	0·00	0·00	0·00
Second moment about origin	0·95	0·98	0·81*	0·93	0·96	0·95	0·95	0·93

Theoretical values of the cumulants of the different distributions are shown in the following table:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
κ_3	0·0	0·0	6·2	2·0	0·5	0·0	0·5	0·0
κ_4	0·0	−1·2	110·9	6·0	0·0	0·5	0·6	3·0
κ_5	0·0	0·0	5154·7	24·0	0·0	0·0	0·0	0·0
κ_6	0·0	6·9	617375	120·0	0·0	0·0	0·0	30·0

The variations in the forms of distributions are fairly well demonstrated by the following frequency table for the whole group of observations:

* Owing to the high value of κ_4 for this distribution, this seemingly extreme value corresponds to a two-tailed probability of greater than 0·5.

Values of the transformations for the Edgeworth Type A distributions

	x_1	x_2*	x_3	x_4	x_5	x_6	x_7	x_8
8·00 to 8·99	.	.	1	.	.	.	.	.
7·00 to 7·99	.	.	2	.	.	.	.	.
6·00 to 6·99	.	.	1	.	.	.	.	.
5·00 to 5·99	.	.	0	3	.	.	.	.
4·00 to 4·99	.	.	5	2	.	.	.	1
3·60 to 3·99	.	.	2	4	.	.	1	3
3·20 to 3·59	.	.	5	7	3	1	3	0
2·80 to 3·19	3	.	3	4	2	3	1	5
2·40 to 2·79	3	.	4	5	11	7	13	8
2·00 to 2·39	14	.	5	23	10	11	8	7
1·60 to 1·99	29	34	20	18	33	27	33	24
1·20 to 1·59	62	117	18	43	60	58	58	40
0·80 to 1·19	100	122	43	59	87	86	79	70
0·40 to 0·79	131	117	70	78	120	136	116	124
0·00 to 0·39	156	108	128	121	145	169	146	216
−0·40 to −0·01	152	120	296	182	154	162	167	210
−0·80 to −0·41	150	110	397	293	154	154	174	152
−1·20 to −0·81	96	141	.	158	117	86	112	55
−1·60 to −1·21	48	95	.	.	66	44	56	32
−2·00 to −1·61	30	36	.	.	38	27	26	21
−2·40 to −2·01	20	.	.	.	.	17	5	12
−2·80 to −2·41	6	.	.	.	.	9	2	12
−3·20 to −2·81	.	.	.	.	.	3	.	5
−3·60 to −3·21	.	.	.	.	.	.	.	1
−4·00 to −3·61	.	.	.	.	.	.	.	.

* The expected numbers in the middle eight groupings here are 115·47. The two end groups have expected numbers 38·12.

REFERENCES

FIELLER, E. C., LEWIS, T. & PEARSON, E. S. (1955). *Random Correlated Normal Deviates.* Cambridge University Press.

KENDALL, M. G. (1949). Tables of autoregressive series. *Biometrika*, **36**, 267–89.

KENDALL, M. G. & BABINGTON SMITH, B. (1939). *Tables of Random Sampling Numbers.* Cambridge University Press.

KENDALL, M. G. & STUART, A. (1958). *The Advanced Theory of Statistics*, vol. 1. Charles Griffin and Co.

RAND CORPORATION (1955). *A Million Random Digits with* 100,000 *Normal Deviates.* The Free Press, Illinois.

TIPPETT, L. H. C. (1927). *Random Sampling Numbers.* Cambridge University Press.

WOLD, H. (1954). *Random Normal Deviates.* Cambridge University Press.

(*a*) $x_5 = \frac{1}{124416}[-9552 + 127{,}225x + 7824x^2 - 40x^3 + 576x^4 - 252x^5]$

x	x_5	x	x_5	x	x_5	x	x_5	x	x_5
−3·2	↑	−1·9	−1·680	−0·6	−0·667	0·7	0·670	2·0	2·227
−3·1		−1·8	−1·625	−0·5	−0·572	0·8	0·782	2·1	2·352
−3·0		−1·7	−1·564	−0·4	−0·476	0·9	0·896	2·2	2·478
−2·9		−1·6	−1·499	−0·3	−0·378	1·0	1·011	2·3	2·603
−2·8		−1·5	−1·429	−0·2	−0·279	1·1	1·127	2·4	2·728
−2·7		−1·4	−1·356	−0·1	−0·178	1·2	1·245	2·5	2·851
−2·6		−1·3	−1·278	0	−0·077	1·3	1·364	2·6	2·972
−2·5	−1·857	−1·2	−1·198	0·1	0·026	1·4	1·484	2·7	3·092
−2·4	−1·849	−1·1	−1·115	0·2	0·130	1·5	1·606	2·8	3·208
−2·3	−1·832	−1·0	−1·029	0·3	0·236	1·6	1·728	2·9	3·322
−2·2	−1·806	−0·9	−0·942	0·4	0·342	1·7	1·852	3·0	3·431
−2·1	−1·771	−0·8	−0·852	0·5	0·450	1·8	1·976	3·1	3·536
−2·0	−1·729	−0·7	−0·760	0·6	0·560	1·9	2·101	3·2	3·635

(*b*) $x_6 = \frac{1}{1536}[1411x + 56x^3 - 3x^5]$

$\pm x$	$\pm x_6$	$\pm x$	$\pm x_6$	$\pm x$	$\pm x_6$	$\pm x$	$\pm x_6$	$\pm x$	$\pm x_6$
0·0	0·000	0·7	0·655	1·4	1·376	2·1	2·187	2·7	2·918
0·1	0·092	0·8	0·753	1·5	1·486	2·2	2·309	2·8	3·036
0·2	0·184	0·9	0·852	1·6	1·599	2·3	2·431	2·9	3·153
0·3	0·277	1·0	0·953	1·7	1·713	2·4	2·553	3·0	3·266
0·4	0·370	1·1	1·056	1·8	1·829	2·5	2·675	3·1	3·375
0·5	0·464	1·2	1·160	1·9	1·947	2·6	2·797	3·2	3·479
0·6	0·559	1·3	1·267	2·0	2·066	.	.	.	.

N.B. Negative values of x give negative values of x_6.

(*c*) $x_7 = \frac{1}{124{,}416}[-12{,}144 + 122{,}878x + 14{,}304x^2 - 1066x^3 - 720x^4 + 261x^5]$

x	x_7	x	x_7	x	x_7	x	x_7	x	x_7
−3·2	−3·111	−1·9	−1·628	−0·6	−0·648	0·7	0·646	2·0	2·244
−3·1	−2·934	−1·8	−1·553	−0·5	−0·562	0·8	0·760	2·1	2·377
−3·0	−2·773	−1·7	−1·480	−0·4	−0·474	0·9	0·876	2·2	2·513
−2·9	−2·625	−1·6	−1·408	−0·3	−0·383	1·0	0·993	2·3	2·651
−2·8	−2·490	−1·5	−1·337	−0·2	−0·290	1·1	1·111	2·4	2·792
−2·7	−2·366	−1·4	−1·265	−0·1	−0·195	1·2	1·232	2·5	2·935
−2·6	−2·251	−1·3	−1·193	0	−0·098	1·3	1·353	2·6	3·082
−2·5	−2·145	−1·2	−1·120	0·1	0·002	1·4	1·476	2·7	3·232
−2·4	−2·046	−1·1	−1·045	0·2	0·104	1·5	1·600	2·8	3·386
−2·3	−1·954	−1·0	−0·970	0·3	0·209	1·6	1·726	2·9	3·545
−2·2	−1·866	−0·9	−0·892	0·4	0·315	1·7	1·853	3·0	3·710
−2·1	−1·784	−0·8	−0·813	0·5	0·424	1·8	1·982	3·1	3·880
−2·0	−1·704	−0·7	−0·731	0·6	0·534	1·9	2·112	3·2	4·056

First fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	−0·84	−1·04	−0·56	−0·78	−0·89	−0·79	−0·84	−0·65
2	1·37	1·44	1·06	1·46	1·45	1·34	1·44	1·25
3	−0·18	−0·25	−0·38	−0·44	−0·26	−0·17	−0·27	−0·11
4	0·35	0·47	−0·11	0·01	0·29	0·32	0·26	0·23
5	2·82	1·72	7·00	5·03	3·23	3·06	3·42	3·78
6	2·12	1·67	3·09	3·07	2·38	2·21	2·40	2·39
7	−0·99	−1·17	−0·59	−0·82	−1·02	−0·94	−0·96	−0·80
8	0·34	0·46	−0·11	0·00	0·28	0·31	0·25	0·22
9	0·64	0·83	0·11	0·34	0·60	0·60	0·58	0·46
10	0·11	0·15	−0·25	−0·22	0·04	0·10	0·01	0·06
11	−0·02	−0·03	−0·31	−0·32	−0·10	−0·02	−0·12	−0·01
12	−0·86	−1·06	−0·57	−0·78	−0·91	−0·81	−0·86	−0·67
13	0·01	0·01	−0·30	−0·30	−0·07	0·01	−0·09	0·01
14	−0·03	−0·04	−0·31	−0·33	−0·11	−0·03	−0·13	−0·02
15	−0·41	−0·55	−0·46	−0·58	−0·49	−0·38	−0·48	−0·27
16	0·62	0·81	0·10	0·32	0·58	0·58	0·56	0·44
17	0·83	1·03	0·30	0·59	0·82	0·78	0·79	0·64
18	−0·78	−0·98	−0·55	−0·75	−0·83	−0·73	−0·80	−0·59
19	0·03	0·04	−0·29	−0·28	−0·05	0·03	−0·07	0·02
20	−2·30	−1·70	−0·72	−0·99	−1·83	−2·43	−1·95	−2·72
21	0·78	0·98	0·25	0·52	0·76	0·73	0·74	0·59
22	0·66	0·85	0·13	0·37	0·63	0·62	0·60	0·48
23	−0·34	−0·46	−0·43	−0·54	−0·42	−0·31	−0·42	−0·22
24	1·16	1·31	0·71	1·10	1·20	1·12	1·18	0·99
25	0·76	0·96	0·23	0·50	0·74	0·71	0·71	0·57
26	−1·00	−1·18	−0·59	−0·83	−1·03	−0·95	−0·97	−0·81
27	−0·32	−0·43	−0·43	−0·53	−0·40	−0·30	−0·40	−0·20
28	0·18	0·25	−0·21	−0·15	0·11	0·17	0·08	0·11
29	0·08	0·11	−0·26	−0·24	0·01	0·07	−0·02	0·05
30	1·98	1·65	2·59	2·73	2·20	2·04	2·22	2·15
31	0·91	1·10	0·39	0·71	0·91	0·86	0·89	0·72
32	−0·32	−0·43	−0·43	−0·53	−0·40	−0·30	−0·40	−0·20
33	−1·77	−1·60	−0·68	−0·96	−1·61	−1·79	−1·53	−1·81
34	0·16	0·22	−0·22	−0·17	0·09	0·15	0·06	0·10
35	0·79	0·99	0·26	0·54	0·77	0·74	0·75	0·60
36	0·20	0·27	−0·20	−0·13	0·13	0·18	0·10	0·12
37	1·31	1·40	0·95	1·35	1·38	1·28	1·37	1·17
38	0·35	0·47	−0·11	0·01	0·29	0·32	0·26	0·23
39	0·74	0·94	0·21	0·47	0·72	0·69	0·69	0·55
40	0·27	0·37	−0·16	−0·07	0·20	0·25	0·18	0·17
41	1·39	1·45	1·09	1·50	1·47	1·37	1·46	1·28
42	−0·14	−0·19	−0·36	−0·41	−0·22	−0·13	−0·23	−0·08
43	−0·76	−0·96	−0·55	−0·75	−0·82	−0·71	−0·78	−0·57
44	0·37	0·50	−0·09	0·03	0·31	0·34	0·28	0·24
45	−2·28	−1·69	−0·72	−0·99	−1·83	−2·41	−1·94	−2·68
46	0·36	0·49	−0·10	0·02	0·30	0·33	0·27	0·23
47	1·37	1·44	1·06	1·46	1·45	1·34	1·44	1·25
48	1·04	1·22	0·55	0·90	1·06	0·99	1·04	0·86
49	−0·67	−0·86	−0·53	−0·71	−0·73	−0·63	−0·71	−0·49
50	0·39	0·53	−0·08	0·05	0·33	0·36	0·30	0·25
$\Sigma(x)$	10·48	11·51	8·42	9·48	10·71	10·17	10·36	9·31
$\Sigma(x^2)$	51·4936	46·6287	76·3842	63·9778	53·0095	52·9419	53·8870	56·9311

Second fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	0·68	0·87	0·15	0·39	0·65	0·64	0·62	0·49
2	−0·77	−0·97	−0·55	−0·75	−0·82	−0·72	−0·79	−0·58
3	−1·39	−1·45	−0·65	−0·91	−1·35	−1·37	−1·26	−1·28
4	−1·91	−1·63	−0·69	−0·97	−1·68	−1·96	−1·64	−2·04
5	0·05	0·07	−0·28	−0·27	−0·03	0·05	−0·05	0·03
6	1·32	1·41	0·97	1·37	1·39	1·29	1·38	1·19
7	−0·67	−0·86	−0·53	−0·71	−0·73	−0·63	−0·71	−0·49
8	−0·81	−1·01	−0·56	−0·77	−0·86	−0·76	−0·82	−0·62
9	0·96	1·15	0·45	0·78	0·97	0·91	0·95	0·77
10	−0·24	−0·33	−0·40	−0·48	−0·32	−0·22	−0·33	−0·15
11	−0·66	−0·85	−0·52	−0·71	−0·72	−0·62	−0·70	−0·48
12	1·08	1·25	0·60	0·97	1·10	1·04	1·09	0·90
13	0·27	0·37	−0·16	−0·07	0·20	0·25	0·18	0·17
14	0·46	0·61	−0·03	0·13	0·41	0·43	0·38	0·31
15	−0·77	−0·97	−0·55	−0·75	−0·82	−0·72	−0·79	−0·58
16	0·00	0·00	−0·30	−0·31	−0·08	0·00	−0·10	0·00
17	1·01	1·19	0·51	0·86	1·02	0·96	1·00	0·82
18	−0·92	−1·11	−0·58	−0·80	−0·96	−0·87	−0·91	−0·73
19	0·98	1·17	0·47	0·81	0·99	0·93	0·97	0·79
20	0·24	0·33	−0·17	−0·10	0·17	0·22	0·15	0·15
21	1·01	1·19	0·51	0·86	1·02	0·96	1·00	0·82
22	1·49	1·50	1·29	1·69	1·59	1·48	1·59	1·41
23	1·16	1·31	0·71	1·10	1·20	1·12	1·18	0·99
24	0·48	0·64	−0·02	0·15	0·43	0·45	0·40	0·33
25	0·40	0·54	−0·07	0·07	0·34	0·37	0·32	0·26
26	−0·12	−0·17	−0·35	−0·40	−0·20	−0·11	−0·21	−0·07
27	−0·85	−1·05	−0·57	−0·78	−0·90	−0·80	−0·85	−0·66
28	−0·46	−0·61	−0·47	−0·61	−0·53	−0·43	−0·53	−0·31
29	0·35	0·47	−0·11	0·01	0·29	0·32	0·26	0·23
30	0·23	0·32	−0·18	−0·11	0·16	0·21	0·14	0·14
31	0·20	0·27	−0·20	−0·13	0·13	0·18	0·10	0·12
32	−0·17	−0·23	−0·37	−0·43	−0·25	−0·16	−0·26	−0·10
33	0·12	0·17	−0·24	−0·21	0·05	0·11	0·02	0·07
34	−1·48	−1·49	−0·66	−0·93	−1·41	−1·46	−1·32	−1·40
35	0·52	0·69	0·02	0·20	0·47	0·48	0·45	0·36
36	1·56	1·53	1·44	1·82	1·68	1·55	1·68	1·51
37	1·04	1·22	0·55	0·90	1·06	0·99	1·04	0·86
38	−0·18	−0·25	−0·38	−0·44	−0·26	−0·17	−0·27	−0·11
39	−0·26	−0·36	−0·41	−0·49	−0·34	−0·24	−0·35	−0·16
40	1·00	1·18	0·49	0·84	1·01	0·95	0·99	0·81
41	−0·51	−0·68	−0·49	−0·64	−0·58	−0·47	−0·57	−0·35
42	0·05	0·07	−0·28	−0·27	−0·03	0·05	−0·05	0·03
43	−0·69	−0·88	−0·53	−0·72	−0·75	−0·65	−0·72	−0·50
44	−0·58	−0·76	−0·50	−0·67	−0·65	−0·54	−0·63	−0·41
45	−0·69	−0·88	−0·53	−0·72	−0·75	−0·65	−0·72	−0·50
46	−1·41	−1·46	−0·65	−0·92	−1·36	−1·39	−1·27	−1·30
47	0·24	0·33	−0·17	−0·10	0·17	0·22	0·15	0·15
48	0·41	0·55	−0·07	0·08	0·35	0·38	0·33	0·27
49	0·15	0·21	−0·23	−0·18	0·08	0·14	0·05	0·09
50	−0·24	−0·33	−0·40	−0·48	−0·32	−0·22	−0·33	−0·15
$\Sigma(x)$	1·68	2·28	−5·69	−3·80	0·23	1·52	0·24	1·10
$\Sigma(x^2)$	32·9330	40·9668	13·6407	26·0324	33·5593	31·0336	31·6978	25·2532

Third fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	0·11	0·15	−0·25	−0·22	0·04	0·10	0·01	0·06
2	−0·89	−1·09	−0·57	−0·79	−0·93	−0·84	−0·88	−0·70
3	0·99	1·17	0·48	0·83	1·00	0·94	0·98	0·80
4	0·50	0·66	0·00	0·18	0·45	0·46	0·42	0·34
5	−0·95	−1·14	−0·58	−0·81	−0·99	−0·90	−0·93	−0·76
6	0·81	1·01	0·28	0·57	0·79	0·76	0·77	0·62
7	−0·28	−0·38	−0·41	−0·51	−0·36	−0·26	−0·36	−0·18
8	−1·64	−1·56	−0·67	−0·95	−1·53	−1·64	−1·44	−1·62
9	−1·44	−1·47	−0·65	−0·92	−1·39	−1·42	−1·29	−1·34
10	0·01	0·01	−0·30	−0·30	−0·07	0·01	−0·09	0·01
11	0·24	0·33	−0·17	−0·10	0·17	0·22	0·15	0·15
12	0·46	0·61	−0·03	0·13	0·41	0·43	0·38	0·31
13	0·02	0·03	−0·29	−0·29	−0·06	0·02	−0·08	0·01
14	−2·09	−1·67	−0·71	−0·98	−1·77	−2·17	−1·78	−2·34
15	2·08	1·67	2·94	2·97	2·33	2·16	2·35	2·32
16	−0·80	−1·00	−0·56	−0·76	−0·85	−0·75	−0·81	−0·61
17	0·27	0·37	−0·16	−0·07	0·20	0·25	0·18	0·17
18	0·21	0·29	−0·19	−0·12	0·14	0·19	0·11	0·13
19	−1·33	−1·41	−0·64	−0·90	−1·30	−1·30	−1·21	−1·20
20	1·70	1·58	1·77	2·11	1·85	1·71	1·85	1·71
21	−0·92	−1·11	−0·58	−0·80	−0·96	−0·87	−0·91	−0·73
22	1·73	1·59	1·85	2·17	1·89	1·75	1·89	1·75
23	0·66	0·85	0·13	0·37	0·63	0·62	0·60	0·48
24	−1·69	−1·57	−0·68	−0·95	−1·56	−1·70	−1·47	−1·19
25	2·13	1·67	3·13	3·10	2·39	2·22	2·42	2·41
26	−2·20	−1·68	−0·71	−0·99	−1·81	−2·31	−1·87	−2·53
27	1·57	1·53	1·46	1·84	1·69	1·57	1·69	1·52
28	−1·85	−1·62	−0·69	−0·97	−1·65	−1·89	−1·59	−1·94
29	−2·24	−1·69	−0·71	−0·99	−1·82	−2·36	−1·90	−2·61
30	−0·15	−0·21	−0·36	−0·42	−0·23	−0·14	−0·24	−0·09
31	−0·45	−0·60	−0·47	−0·60	−0·52	−0·42	−0·52	−0·30
32	−0·03	−0·04	−0·31	−0·33	−0·11	−0·03	−0·13	−0·02
33	1·01	1·19	0·51	0·86	1·02	0·96	1·00	0·82
34	1·19	1·33	0·76	1·15	1·23	1·15	1·22	1·03
35	−1·03	−1·21	−0·60	−0·84	−1·05	−0·98	−0·99	−0·84
36	0·35	0·47	−0·11	0·01	0·29	0·32	0·26	0·23
37	−0·82	−1·02	−0·56	−0·77	−0·87	−0·77	−0·83	−0·63
38	1·76	1·60	1·93	2·24	1·93	1·78	1·93	1·80
39	−0·68	−0·87	−0·53	−0·71	−0·74	−0·64	−0·71	−0·49
40	−0·88	−1·08	−0·57	−0·79	−0·92	−0·83	−0·88	−0·69
41	−0·14	−0·19	−0·36	−0·41	−0·22	−0·13	−0·23	−0·08
42	0·77	0·97	0·24	0·51	0·75	0·72	0·73	0·58
43	−1·61	−1·55	−0·67	−0·94	−1·51	−1·61	−1·42	−1·58
44	−0·67	−0·86	−0·53	−0·71	−0·73	−0·63	−0·71	−0·49
45	0·58	0·76	0·06	0·27	0·54	0·54	0·51	0·41
46	−1·06	−1·23	−0·60	−0·84	−1·08	−1·01	−1·02	−0·88
47	0·51	0·68	0·01	0·19	0·46	0·47	0·44	0·35
48	−0·24	−0·33	−0·40	−0·48	−0·32	−0·22	−0·33	−0·15
49	−1·60	−1·54	−0·67	−0·94	−1·50	−1·60	−1·41	−1·56
50	0·29	0·40	−0·14	−0·05	0·23	0·27	0·20	0·18
$\Sigma(x)$	−7·73	−7·20	−0·88	−1·75	−6·42	−7·80	−5·94	−7·36
$\Sigma(x^2)$	67·1551	62·3312	41·2024	56·4089	65·3910	67·9222	63·2752	66·3356

Fourth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	1·05	1·22	0·56	0·92	1·07	1·00	1·05	0·87
2	0·61	0·79	0·09	0·31	0·57	0·57	0·55	0·43
3	−0·01	−0·01	−0·30	−0·31	−0·09	−0·01	−0·11	−0·01
4	−0·92	−1·11	−0·58	−0·80	−0·96	−0·87	−0·91	−0·73
5	1·17	1·31	0·73	1·11	1·21	1·13	1·20	1·00
6	−1·91	−1·63	−0·69	−0·97	−1·68	−1·96	−1·64	−2·04
7	−0·48	−0·64	−0·48	−0·62	−0·55	−0·45	−0·54	−0·33
8	−0·21	−0·29	−0·39	−0·46	−0·29	−0·19	−0·30	−0·13
9	−2·47	−1·71	−0·72	−0·99	−1·85	−2·64	−2·12	−3·04
10	−0·28	−0·38	−0·41	−0·51	−0·36	−0·26	−0·36	−0·18
11	−0·43	−0·58	−0·46	−0·59	−0·50	−0·40	−0·50	−0·29
12	0·08	0·11	−0·26	−0·24	0·01	0·07	−0·02	0·05
13	−0·60	−0·78	−0·51	−0·68	−0·67	−0·56	−0·65	−0·42
14	−1·60	−1·54	−0·67	−0·94	−1·50	−1·60	−1·41	−1·56
15	−0·98	−1·17	−0·59	−0·82	−1·01	−0·93	−0·95	−0·79
16	−0·62	−0·81	−0·51	−0·69	−0·69	−0·58	−0·66	−0·44
17	1·47	1·49	1·25	1·65	1·57	1·45	1·56	1·38
18	−0·60	−0·78	−0·51	−0·68	−0·67	−0·56	−0·65	−0·42
19	0·66	0·85	0·13	0·37	0·63	0·62	0·60	0·48
20	−0·52	−0·69	−0·49	−0·64	−0·59	−0·48	−0·58	−0·36
21	1·82	1·61	2·09	2·37	2·00	1·85	2·01	1·89
22	−0·27	−0·37	−0·41	−0·50	−0·35	−0·25	−0·36	−0·17
23	0·59	0·77	0·07	0·28	0·55	0·55	0·52	0·42
24	0·07	0·10	−0·27	−0·25	0·00	0·06	−0·03	0·04
25	−0·25	−0·34	−0·40	−0·49	−0·33	−0·23	−0·34	−0·16
26	0·75	0·95	0·22	0·48	0·73	0·70	0·70	0·56
27	0·25	0·34	−0·17	−0·09	0·18	0·23	0·16	0·16
28	2·24	1·69	3·58	3·38	2·53	2·36	2·57	2·61
29	−0·27	−0·37	−0·41	−0·50	−0·35	−0·25	−0·36	−0·17
30	−0·22	−0·30	−0·39	−0·47	−0·30	−0·20	−0·31	−0·14
31	0·06	0·08	−0·27	−0·26	−0·02	0·06	−0·04	0·03
32	−0·75	−0·95	−0·54	−0·74	−0·81	−0·70	−0·77	−0·56
33	−2·27	−1·69	−0·72	−0·99	−1·82	−2·39	−1·93	−2·66
34	−0·23	−0·32	−0·40	−0·47	−0·31	−0·21	−0·32	−0·14
35	−1·18	−1·32	−0·62	−0·87	−1·18	−1·14	−1·11	−1·02
36	−0·62	−0·81	−0·51	−0·69	−0·69	−0·58	−0·66	−0·44
37	0·47	0·63	−0·02	0·14	0·42	0·44	0·39	0·32
38	0·92	1·11	0·40	0·72	0·92	0·87	0·90	0·73
39	−0·32	−0·43	−0·43	−0·53	−0·40	−0·30	−0·40	−0·20
40	−0·23	−0·32	−0·40	−0·47	−0·31	−0·21	−0·32	−0·14
41	−0·89	−1·09	−0·57	−0·79	−0·93	−0·84	−0·88	−0·70
42	1·33	1·41	0·99	1·39	1·40	1·30	1·39	1·20
43	0·33	0·45	−0·12	−0·01	0·27	0·30	0·24	0·21
44	0·68	0·87	0·15	0·39	0·65	0·64	0·62	0·49
45	−0·38	−0·51	−0·45	−0·57	−0·46	−0·35	−0·46	−0·25
46	−0·89	−1·09	−0·57	−0·79	−0·93	−0·84	−0·88	−0·70
47	−0·28	−0·38	−0·41	−0·51	−0·36	−0·26	−0·36	−0·28
48	0·19	0·26	−0·20	−0·14	0·12	0·17	0·09	0·12
49	0·84	1·04	0·31	0·61	0·83	0·79	0·81	0·65
50	−0·03	−0·04	−0·31	−0·33	−0·11	−0·03	−0·13	−0·02
$\Sigma(x)$	−5·13	−5·37	−5·59	−6·28	−5·41	−5·11	−5·70	−4·85
$\Sigma(x^2)$	44·7007	43·2703	28·9523	39·2082	42·8955	45·2211	43·0940	45·6167

Fifth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	0·09	0·12	−0·26	−0·23	0·02	0·08	−0·01	0·05
2	−2·54	−1·71	−0·73	−0·99	−1·86	−2·72	−2·19	−3·19
3	−0·22	−0·30	−0·39	−0·47	−0·30	−0·20	−0·31	−0·14
4	−0·30	−0·41	−0·42	−0·52	−0·38	−0·28	−0·38	−0·19
5	1·22	1·35	0·80	1·20	1·27	1·18	1·26	1·06
6	−1·23	−1·35	−0·63	−0·88	−1·22	−1·19	−1·14	−1·07
7	−0·12	−0·17	−0·35	−0·40	−0·20	−0·11	−0·21	−0·07
8	−0·19	−0·26	−0·38	−0·45	−0·27	−0·17	−0·28	−0·12
9	−1·69	−1·57	−0·68	−0·95	−1·56	−1·70	−1·47	−1·69
10	0·04	0·06	−0·28	−0·27	−0·04	0·04	−0·06	0·02
11	1·09	1·25	0·61	0·98	1·12	1·05	1·10	0·91
12	1·03	1·21	0·53	0·89	1·05	0·98	1·03	0·84
13	−0·61	−0·79	−0·51	−0·68	−0·68	−0·57	−0·66	−0·43
14	−0·56	−0·74	−0·50	−0·66	−0·63	−0·52	−0·61	−0·39
15	0·36	0·49	−0·10	0·02	0·30	0·33	0·27	0·23
16	0·35	0·47	−0·11	0·01	0·29	0·32	0·26	0·23
17	0·29	0·40	−0·14	−0·05	0·23	0·27	0·20	0·18
18	−0·59	−0·77	−0·51	−0·67	−0·66	−0·55	−0·64	−0·42
19	−0·57	−0·75	−0·50	−0·67	−0·64	−0·53	−0·62	−0·40
20	−0·15	−0·21	−0·36	−0·42	−0·23	−0·14	−0·24	−0·09
21	−0·67	−0·86	−0·53	−0·71	−0·73	−0·63	−0·71	−0·49
22	−0·87	−1·07	−0·57	−0·79	−0·92	−0·82	−0·87	−0·68
23	0·35	0·47	−0·11	0·01	0·29	0·32	0·26	0·23
24	0·70	0·89	0·17	0·42	0·67	0·66	0·65	0·51
25	−0·15	−0·21	−0·36	−0·42	−0·23	−0·14	−0·24	−0·09
26	1·30	1·40	0·93	1·34	1·36	1·27	1·35	1·16
27	0·75	0·95	0·22	0·48	0·73	0·70	0·70	0·56
28	1·08	1·25	0·60	0·97	1·10	1·04	1·09	0·90
29	0·67	0·86	0·14	0·38	0·64	0·63	0·61	0·49
30	−0·56	−0·74	−0·50	−0·66	−0·63	−0·52	−0·61	−0·39
31	−1·10	−1·26	−0·61	−0·85	−1·12	−1·06	−1·05	−0·92
32	0·93	1·12	0·41	0·74	0·93	0·88	0·91	0·74
33	0·55	0·72	0·04	0·23	0·51	0·51	0·48	0·38
34	0·59	0·77	0·07	0·28	0·55	0·55	0·52	0·42
35	−0·85	−1·05	−0·57	−0·78	−0·90	−0·80	−0·85	−0·66
36	−0·15	−0·21	−0·36	−0·42	−0·23	−0·14	−0·24	−0·09
37	0·81	1·01	0·28	0·57	0·79	0·76	0·77	0·62
38	−1·21	−1·34	−0·62	−0·88	−1·21	−1·17	−1·13	−1·05
39	−0·44	−0·59	−0·46	−0·60	−0·51	−0·41	−0·51	−0·30
40	−0·47	−0·63	−0·47	−0·62	−0·54	−0·44	−0·54	−0·32
41	0·76	0·96	0·23	0·50	0·74	0·71	0·71	0·57
42	−0·58	−0·76	−0·50	−0·67	−0·65	−0·54	−0·63	−0·41
43	−0·13	−0·18	−0·36	−0·41	−0·21	−0·12	−0·22	−0·08
44	0·17	0·23	−0·21	−0·16	0·10	0·16	0·07	0·10
45	−1·65	−1·56	−0·67	−0·95	−1·53	−1·66	−1·44	−1·64
46	−0·28	−0·38	−0·41	−0·51	−0·36	−0·26	−0·36	−0·17
47	0·32	0·43	−0·13	−0·02	0·26	0·30	0·23	0·20
48	−2·10	−1·67	−0·71	−0·98	−1·77	−2·19	−1·78	−2·35
49	−0·07	−0·10	−0·33	−0·36	−0·15	−0·06	−0·17	−0·04
50	0·83	1·03	0·30	0·59	0·82	0·78	0·79	0·64
$\Sigma(x)$	−5·77	−4·20	−10·00	−9·49	−6·59	−6·12	−6·91	−6·84
$\Sigma(x^2)$	37·4683	40·9470	10·6474	21·3559	33·6241	36·8738	33·0399	34·6564

Sixth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	0·83	1·03	0·30	0·59	0·82	0·78	0·79	0·64
2	1·55	1·52	1·42	1·80	1·67	1·54	1·66	1·49
3	0·36	0·49	−0·10	0·02	0·30	0·33	0·27	0·23
4	0·36	0·49	−0·10	0·02	0·30	0·33	0·27	0·23
5	0·14	0·19	−0·23	−0·19	0·07	0·13	0·04	0·08
6	−0·17	−0·23	−0·37	−0·43	−0·25	−0·16	−0·26	−0·10
7	−2·21	−1·68	−0·71	−0·99	−1·81	−2·32	−1·87	−2·55
8	0·66	0·85	0·13	0·37	0·63	0·62	0·60	0·48
9	−0·45	−0·60	−0·47	−0·60	−0·52	−0·42	−0·52	−0·30
10	1·28	1·38	0·90	1·30	1·34	1·25	1·33	1·14
11	−0·41	−0·55	−0·46	−0·58	−0·49	−0·38	−0·48	−0·27
12	−1·73	−1·59	−0·68	−0·96	−1·58	−1·75	−1·50	−1·75
13	−1·01	−1·19	−0·59	−0·83	−1·04	−0·96	−0·98	−0·82
14	−1·38	−1·44	−0·65	−0·91	−1·34	−1·35	−1·25	−1·26
15	−0·36	−0·49	−0·44	−0·55	−0·44	−0·33	−0·44	−0·23
16	0·31	0·42	−0·13	−0·03	0·25	0·29	0·22	0·19
17	−2·79	−1·72	−0·73	−1·00	−1·86*	−3·02	−2·48	−3·72
18	0·14	0·19	−0·23	−0·19	0·07	0·13	0·04	0·08
19	−0·68	−0·87	−0·53	−0·71	−0·74	−0·64	−0·71	−0·49
20	1·13	1·28	0·67	1·05	1·16	1·09	1·15	0·96
21	−1·30	−1·40	−0·64	−0·90	−1·28	−1·27	−1·19	−1·16
22	−0·64	−0·83	−0·52	−0·70	−0·70	−0·60	−0·68	−0·46
23	−1·63	−1·55	−0·67	−0·95	−1·52	−1·63	−1·43	−1·61
24	−0·49	−0·65	−0·48	−0·63	−0·56	−0·45	−0·55	−0·33
25	−0·25	−0·34	−0·40	−0·49	−0·33	−0·23	−0·34	−0·16
26	−0·64	−0·83	−0·52	−0·70	−0·70	−0·60	−0·68	−0·46
27	−0·61	−0·79	−0·51	−0·68	−0·68	−0·57	−0·66	−0·43
28	0·24	0·33	−0·17	−0·10	0·17	0·22	0·15	0·15
29	−1·01	−1·19	−0·59	−0·83	−1·04	−0·96	−0·98	−0·82
30	−0·62	−0·81	−0·51	−0·69	−0·69	−0·58	−0·66	−0·44
31	−0·23	−0·32	−0·40	−0·47	−0·31	−0·21	−0·32	−0·14
32	0·42	0·56	−0·06	0·09	0·36	0·39	0·34	0·28
33	0·32	0·43	−0·13	−0·02	0·26	0·30	0·23	0·20
34	0·41	0·55	−0·07	0·08	0·35	0·38	0·33	0·27
35	1·67	1·57	1·70	2·05	1·81	1·68	1·81	1·66
36	−0·19	−0·26	−0·38	−0·45	−0·27	−0·17	−0·28	−0·12
37	0·26	0·36	−0·16	−0·08	0·19	0·24	0·17	0·16
38	−0·48	−0·64	−0·48	−0·62	−0·55	−0·45	−0·54	−0·33
39	0·01	0·01	−0·30	−0·30	−0·07	0·01	−0·09	0·01
40	−1·61	−1·55	−0·67	−0·94	−1·51	−1·61	−1·42	−1·58
41	0·14	0·19	−0·23	−0·19	0·07	0·13	0·04	0·08
42	−0·76	−0·96	−0·55	−0·75	−0·82	−0·71	−0·78	−0·57
43	0·30	0·41	−0·14	−0·04	0·24	0·28	0·21	0·19
44	−0·54	−0·71	−0·49	−0·65	−0·61	−0·50	−0·60	−0·37
45	−0·86	−1·06	−0·57	−0·78	−0·91	−0·81	−0·86	−0·67
46	−0·75	−0·95	−0·54	−0·74	−0·81	−0·70	−0·77	−0·56
47	0·32	0·43	−0·13	−0·02	0·26	0·30	0·23	0·20
48	−0·69	−0·88	−0·53	−0·72	−0·75	−0·65	−0·72	−0·50
49	−0·06	−0·08	−0·33	−0·35	−0·14	−0·06	−0·16	−0·03
50	0·67	0·86	0·14	0·38	0·64	0·63	0·61	0·49
$\Sigma(x)$	−13·03	−12·62	−12·33	−14·01	−13·36	−13·04	−13·71	−13·02
$\Sigma(x^2)$	43·2437	42·6990	15·1983	26·6345	38·1144	43·5442	38·7677	44·3438

Seventh fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	1·35	1·43	1·02	1·42	1·42	1·32	1·41	1·22
2	−0·92	−1·11	−0·58	−0·80	−0·96	−0·87	−0·91	−0·73
3	1·27	1·38	0·88	1·28	1·33	1·23	1·32	1·12
4	1·28	1·38	0·90	1·30	1·34	1·25	1·33	1·14
5	−2·13	−1·67	−0·70	−0·98	−1·78	−2·22	−1·81	−2·41
6	−0·38	−0·51	−0·45	−0·57	−0·46	−0·35	−0·46	−0·25
7	−0·10	−0·14	−0·34	−0·38	−0·18	−0·09	−0·20	−0·06
8	0·12	0·17	−0·24	−0·21	0·05	0·11	0·02	0·07
9	2·31	1·70	3·90	3·57	2·62	2·44	2·67	2·74
10	0·47	0·63	−0·02	0·14	0·42	0·44	0·39	0·32
11	0·42	0·56	−0·06	0·09	0·36	0·39	0·34	0·28
12	−1·07	−1·24	−0·60	−0·85	−1·09	−1·03	−1·02	−0·89
13	0·22	0·30	−0·19	−0·12	0·15	0·20	0·13	0·14
14	−1·33	−1·41	−0·64	−0·90	−1·30	−1·30	−1·21	−1·20
15	−2·24	−1·69	−0·71	−0·99	−1·82	−2·36	−1·90	−2·61
16	−0·20	−0·27	−0·38	−0·45	−0·28	−0·18	−0·29	−0·12
17	0·20	0·27	−0·20	−0·13	0·13	0·18	0·10	0·12
18	0·09	0·12	−0·26	−0·23	0·02	0·08	−0·01	0·05
19	0·02	0·03	−0·29	−0·29	−0·06	0·02	−0·08	0·01
20	−0·23	−0·32	−0·40	−0·47	−0·31	−0·21	−0·32	−0·14
21	−1·07	−1·24	−0·60	−0·85	−1·09	−1·03	−1·02	−0·89
22	1·43	1·47	1·17	1·57	1·52	1·41	1·51	1·33
23	0·22	0·30	−0·19	−0·12	0·15	0·20	0·13	0·14
24	−1·24	−1·36	−0·63	−0·89	−1·23	−1·20	−1·15	−1·09
25	0·91	1·10	0·39	0·71	0·91	0·86	0·89	0·72
26	−1·88	−1·63	−0·69	−0·97	−1·67	−1·92	−1·61	−1·99
27	0·08	0·11	−0·26	−0·24	0·01	0·07	−0·02	0·05
28	2·19	1·68	3·37	3·25	2·47	2·30	2·50	2·51
29	0·17	0·23	−0·21	−0·16	0·10	0·16	0·07	0·10
30	−0·61	−0·79	−0·51	−0·68	−0·68	−0·57	−0·66	−0·43
31	−0·12	−0·17	−0·35	−0·40	−0·20	−0·11	−0·21	−0·07
32	0·70	0·89	0·17	0·42	0·67	0·66	0·65	0·51
33	−1·72	−1·58	−0·68	−0·96	−1·58	−1·74	−1·49	−1·74
34	−1·46	−1·48	−0·66	−0·93	−1·40	−1·44	−1·31	−1·37
35	−0·36	−0·49	−0·44	−0·55	−0·44	−0·33	−0·44	−0·23
36	−1·39	−1·45	−0·65	−0·91	−1·35	−1·37	−1·26	−1·28
37	0·01	0·01	−0·30	−0·30	−0·07	0·01	−0·09	0·01
38	0·52	0·69	0·02	0·20	0·47	0·48	0·45	0·36
39	−0·47	−0·63	−0·47	−0·62	−0·54	−0·44	−0·54	−0·32
40	−0·01	−0·01	−0·30	−0·31	−0·09	−0·01	−0·11	−0·01
41	−0·49	−0·65	−0·48	−0·63	−0·56	−0·45	−0·55	−0·33
42	0·19	0·26	−0·20	−0·14	0·12	0·17	0·09	0·12
43	−0·05	−0·07	−0·32	−0·35	−0·13	−0·05	−0·15	−0·03
44	1·07	1·24	0·59	0·95	1·09	1·03	1·08	0·89
45	−1·39	−1·45	−0·65	−0·91	−1·35	−1·37	−1·26	−1·28
46	0·96	1·15	0·45	0·78	0·97	0·91	0·95	0·77
47	1·32	1·41	0·97	1·37	1·39	1·29	1·38	1·19
48	−0·98	−1·17	−0·59	−0·82	−1·01	−0·93	−0·95	−0·79
49	−0·25	−0·34	−0·40	−0·49	−0·33	−0·23	−0·34	−0·16
50	−1·20	−1·33	−0·62	−0·88	−1·20	−1·16	−1·12	−1·54
$\Sigma(x)$	−5·77	−5·69	−2·43	−3·43	−5·45	−5·75	−5·08	−6·05
$\Sigma(x^2)$	55·4525	52·9583	40·6117	50·1883	55·0049	55·9079	53·1864	56·6803

Eighth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	1·01	1·19	0·51	0·86	1·02	0·96	1·00	0·82
2	−0·20	−0·27	−0·38	−0·45	−0·28	−0·18	−0·29	−0·12
3	0·45	0·60	−0·04	0·12	0·40	0·42	0·37	0·30
4	−0·69	−0·88	−0·53	−0·72	−0·75	−0·65	−0·72	−0·50
5	−0·76	−0·96	−0·55	−0·75	−0·82	−0·71	−0·78	−0·57
6	−0·03	−0·04	−0·31	−0·33	−0·11	−0·03	−0·13	−0·02
7	0·40	0·54	−0·07	0·07	0·34	0·37	0·32	0·26
8	0·39	0·53	−0·08	0·05	0·33	0·36	0·30	0·25
9	1·39	1·45	1·09	1·50	1·47	1·37	1·46	1·28
10	0·77	0·97	0·24	0·51	0·75	0·72	0·73	0·58
11	−0·37	−0·50	−0·44	−0·56	−0·45	−0·34	−0·45	−0·24
12	0·35	0·47	−0·11	0·01	0·29	0·32	0·26	0·23
13	−1·25	−1·37	−0·63	−0·89	−1·24	−1·21	−1·16	−1·10
14	−0·70	−0·89	−0·53	−0·72	−0·76	−0·66	−0·73	−0·51
15	1·41	1·46	1·13	1·53	1·50	1·39	1·49	1·30
16	−0·04	−0·06	−0·32	−0·34	−0·12	−0·04	−0·14	−0·02
17	0·42	0·56	−0·06	0·09	0·36	0·39	0·34	0·28
18	−0·09	−0·12	−0·34	−0·38	−0·17	−0·08	−0·19	−0·05
19	−0·59	−0·77	−0·51	−0·67	−0·66	−0·55	−0·64	−0·42
20	0·77	0·97	0·24	0·51	0·75	0·72	0·73	0·58
21	−0·69	−0·88	−0·53	−0·72	−0·75	−0·65	−0·72	−0·50
22	0·14	0·19	−0·23	−0·19	0·07	0·13	0·04	0·08
23	−0·34	−0·46	−0·43	−0·54	−0·42	−0·31	−0·42	−0·22
24	−0·51	−0·68	−0·49	−0·64	−0·58	−0·47	−0·57	−0·35
25	0·54	0·71	0·03	0·22	0·49	0·50	0·47	0·37
26	1·35	1·43	1·02	1·42	1·42	1·32	1·41	1·22
27	0·06	0·08	−0·27	−0·26	−0·02	0·06	−0·04	0·03
28	0·66	0·85	0·13	0·37	0·63	0·62	0·60	0·48
29	−0·35	−0·47	−0·44	−0·55	−0·43	−0·32	−0·43	−0·23
30	−0·53	−0·70	−0·49	−0·65	−0·60	−0·49	−0·59	−0·37
31	−0·56	−0·74	−0·50	−0·66	−0·63	−0·52	−0·61	−0·39
32	−0·02	−0·03	−0·31	−0·32	−0·10	−0·02	−0·12	−0·01
33	0·66	0·85	0·13	0·37	0·63	0·62	0·60	0·48
34	0·10	0·14	−0·25	−0·22	0·03	0·09	0·00	0·06
35	0·52	0·69	0·02	0·20	0·47	0·48	0·45	0·36
36	−0·46	−0·61	−0·47	−0·61	−0·53	−0·43	−0·53	−0·31
37	0·47	0·63	−0·02	0·14	0·42	0·44	0·39	0·32
38	0·40	0·54	−0·07	0·07	0·34	0·37	0·32	0·26
39	−1·15	−1·30	−0·62	−0·87	−1·16	−1·11	−1·08	−0·98
40	−0·29	−0·40	−0·42	−0·51	−0·37	−0·27	−0·37	−0·18
41	0·82	1·02	0·29	0·58	0·81	0·77	0·78	0·63
42	0·23	0·32	−0·18	−0·11	0·16	0·21	0·14	0·14
43	−0·87	−1·07	−0·57	−0·79	−0·92	−0·82	−0·87	−0·68
44	−0·12	−0·17	−0·35	−0·40	−0·20	−0·11	−0·21	−0·07
45	0·64	0·83	0·11	0·34	0·60	0·60	0·58	0·46
46	−0·03	−0·04	−0·31	−0·33	−0·11	−0·03	−0·13	−0·02
47	0·17	0·23	−0·21	−0·16	0·10	0·16	0·07	0·10
48	1·49	1·50	1·29	1·69	1·59	1·48	1·59	1·41
49	−0·51	−0·68	−0·49	−0·64	−0·58	−0·47	−0·57	−0·35
50	0·62	0·81	0·10	0·32	0·58	0·58	0·56	0·44
$\Sigma(x)$	5·08	5·47	−6·22	−4·01	2·79	4·98	2·51	4·51
$\Sigma(x^2)$	22·2392	30·8785	11·2092	20·7837	23·8075	20·3996	22·5705	14·9801

Ninth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	−0·54	−0·71	−0·49	−0·65	−0·61	−0·50	−0·60	−0·37
2	1·11	1·27	0·64	1·01	1·14	1·07	1·12	0·94
3	0·59	0·77	0·07	0·28	0·55	0·55	0·52	0·42
4	−1·64	−1·56	−0·67	−0·95	−1·53	−1·64	−1·44	−1·62
5	0·24	0·33	−0·17	−0·10	0·17	0·22	0·15	0·15
6	−1·65	−1·56	−0·67	−0·95	−1·53	−1·66	−1·44	−1·64
7	−0·83	−1·03	−0·56	−0·77	−0·88	−0·78	−0·84	−0·64
8	1·06	1·23	0·57	0·93	1·08	1·01	1·06	0·88
9	−0·17	−0·23	−0·37	−0·43	−0·25	−0·16	−0·26	−0·10
10	−1·31	−1·40	−0·64	−0·90	−1·29	−1·28	−1·20	−1·17
11	−0·56	−0·74	−0·50	−0·66	−0·63	−0·52	−0·61	−0·39
12	0·66	0·85	0·13	0·37	0·63	0·62	0·60	0·48
13	1·43	1·47	1·17	1·57	1·52	1·41	1·51	1·33
14	−0·18	−0·25	−0·38	−0·44	−0·26	−0·17	−0·27	−0·11
15	−0·07	−0·10	−0·33	−0·36	−0·15	−0·06	−0·17	−0·04
16	−0·41	−0·55	−0·46	−0·58	−0·49	−0·38	−0·48	−0·27
17	0·51	0·68	0·01	0·19	0·46	0·47	0·44	0·35
18	0·97	1·16	0·46	0·80	0·98	0·92	0·96	0·78
19	−1·20	−1·33	−0·62	−0·88	−1·20	−1·16	−1·12	−1·54
20	−0·10	−0·14	−0·34	−0·38	−0·18	−0·09	−0·20	−0·06
21	1·53	1·51	1·37	1·76	1·64	1·52	1·64	1·46
22	−0·68	−0·87	−0·53	−0·71	−0·74	−0·64	−0·71	−0·49
23	1·70	1·58	1·77	2·11	1·85	1·71	1·85	1·71
24	1·40	1·45	1·11	1·52	1·48	1·38	1·48	1·29
25	0·17	0·23	−0·21	−0·16	0·10	0·16	0·07	0·10
26	−1·98	−1·65	−0·70	−0·98	−1·72	−2·04	−1·69	−2·15
27	−0·61	−0·79	−0·51	−0·68	−0·68	−0·57	−0·66	−0·43
28	1·03	1·21	0·53	0·89	1·05	0·98	1·03	0·84
29	0·66	0·85	0·13	0·37	0·63	0·62	0·60	0·48
30	−0·62	−0·81	−0·51	−0·69	−0·69	−0·58	−0·66	−0·44
31	−0·03	−0·04	−0·31	−0·33	−0·11	−0·03	−0·13	−0·02
32	0·15	0·21	−0·23	−0·18	0·08	0·14	0·05	0·09
33	−1·86	−1·62	−0·69	−0·97	−1·66	−1·90	−1·60	−1·96
34	−1·55	−1·52	−0·66	−0·94	−1·46	−1·54	−1·37	−1·49
35	0·34	0·46	−0·11	0·00	0·28	0·31	0·25	0·22
36	−0·63	−0·82	−0·52	−0·69	−0·69	−0·59	−0·67	−0·45
37	0·30	0·41	−0·14	−0·04	0·24	0·28	0·21	0·19
38	0·31	0·42	−0·13	−0·03	0·25	0·29	0·22	0·19
39	1·13	1·28	0·67	1·05	1·16	1·09	1·15	0·96
40	0·68	0·87	0·15	0·39	0·65	0·64	0·62	0·49
41	−0·01	−0·01	−0·30	−0·31	−0·09	−0·01	−0·11	−0·01
42	−0·45	−0·60	−0·47	−0·60	−0·52	−0·42	−0·52	−0·30
43	0·03	0·04	−0·29	−0·28	−0·05	0·03	−0·07	0·02
44	0·03	0·04	−0·29	−0·28	−0·05	0·03	−0·07	0·02
45	−0·64	−0·83	−0·52	−0·70	−0·70	−0·60	−0·68	−0·46
46	0·34	0·46	−0·11	0·00	0·28	0·31	0·25	0·22
47	0·01	0·01	−0·30	−0·30	−0·07	0·01	−0·09	0·01
48	0·49	0·65	−0·01	0·16	0·44	0·45	0·41	0·33
49	−0·63	−0·82	−0·52	−0·69	−0·69	−0·59	−0·67	−0·45
50	−2·07	−1·67	−0·70	−0·98	−1·76	−2·15	−1·76	−2·30
$\Sigma(x)$	−3·55	−2·21	−6·18	−5·19	−4·02	−3·84	−3·90	−4·95
$\Sigma(x^2)$	44·3801	47·3053	16·9168	30·6523	42·6600	43·4802	40·4120	40·8547

Tenth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	0·28	0·38	−0·15	−0·06	0·21	0·26	0·19	0·17
2	−1·52	−1·51	−0·66	−0·93	−1·44	−1·51	−1·35	−1·45
3	1·52	1·51	1·35	1·74	1·63	1·51	1·63	1·45
4	0·42	0·56	−0·06	0·09	0·36	0·39	0·34	0·28
5	−1·07	−1·24	−0·60	−0·85	−1·09	−1·03	−1·02	−0·89
6	−0·35	−0·47	−0·44	−0·55	−0·43	−0·32	−0·43	−0·23
7	−0·37	−0·50	−0·44	−0·56	−0·45	−0·34	−0·45	−0·24
8	−0·38	−0·51	−0·45	−0·57	−0·46	−0·35	−0·46	−0·25
9	0·09	0·12	−0·26	−0·23	0·02	0·08	−0·01	0·05
10	−0·38	−0·51	−0·45	−0·57	−0·46	−0·35	−0·46	0·25
11	0·38	0·51	−0·09	0·04	0·32	0·35	0·29	0·25
12	0·88	1·08	0·35	0·66	0·87	0·83	0·85	0·69
13	−1·92	−1·64	−0·69	−0·97	−1·69	−1·97	−1·64	−2·05
14	1·11	1·27	0·64	1·01	1·14	1·07	1·12	0·94
15	−1·59	−1·54	−0·67	−0·94	−1·49	−1·59	−1·40	−1·55
16	−0·33	−0·45	−0·43	−0·54	−0·41	−0·30	−0·41	−0·21
17	−0·63	−0·82	−0·52	−0·69	−0·69	−0·59	−0·67	−0·45
18	0·96	1·15	0·45	0·78	0·97	0·91	0·95	0·77
19	1·05	1·22	0·56	0·92	1·07	1·00	1·05	0·87
20	−0·48	−0·64	−0·48	−0·62	−0·55	−0·45	−0·54	−0·33
21	−0·60	−0·78	−0·51	−0·68	−0·67	−0·56	−0·65	−0·42
22	0·69	0·88	0·16	0·41	0·66	0·65	0·63	0·50
23	−0·46	−0·61	−0·47	−0·61	−0·53	−0·43	−0·53	−0·31
24	−0·87	−1·07	−0·57	−0·79	−0·92	−0·82	−0·87	−0·68
25	−0·45	−0·60	−0·47	−0·60	−0·52	−0·42	−0·52	−0·30
26	−0·14	−0·19	−0·36	−0·41	−0·22	−0·13	−0·23	−0·08
27	−1·02	−1·20	−0·60	−0·83	−1·05	−0·97	−0·99	−0·83
28	−0·58	−0·76	−0·50	−0·67	−0·65	−0·54	−0·63	−0·41
29	1·38	1·44	1·08	1·48	1·46	1·35	1·45	1·26
30	1·11	1·27	0·64	1·01	1·14	1·07	1·12	0·94
31	−1·49	−1·50	−0·66	−0·93	−1·42	−1·48	−1·33	−1·41
32	−0·14	−0·19	−0·36	−0·41	−0·22	−0·13	−0·23	−0·08
33	0·50	0·66	0·00	0·18	0·45	0·46	0·42	0·34
34	−0·35	−0·47	−0·44	−0·55	−0·43	−0·32	−0·43	−0·23
35	0·94	1·13	0·42	0·75	0·94	0·89	0·92	0·75
36	0·50	0·66	0·00	0·18	0·45	0·46	0·42	0·34
37	0·83	1·03	0·30	0·59	0·82	0·78	0·79	0·64
38	0·04	0·06	−0·28	−0·27	−0·04	0·04	−0·06	0·02
39	−0·12	−0·17	−0·35	−0·40	−0·20	−0·11	−0·21	−0·07
40	0·30	0·41	−0·14	−0·04	0·24	0·28	0·21	0·19
41	−1·45	−1·48	−0·65	−0·92	−1·39	−1·43	−1·30	−1·36
42	−0·23	−0·32	−0·40	−0·47	−0·31	−0·21	−0·32	−0·14
43	0·77	0·97	0·24	0·51	0·75	0·72	0·73	0·58
44	−0·37	−0·50	−0·44	−0·56	−0·45	−0·34	−0·45	−0·24
45	1·26	1·37	0·87	1·27	1·32	1·22	1·30	1·11
46	−0·62	−0·81	−0·51	−0·69	−0·69	−0·58	−0·66	−0·44
47	0·00	0·00	−0·30	−0·31	−0·08	0·00	−0·10	0·00
48	0·83	1·03	0·30	0·59	0·82	0·78	0·79	0·64
49	0·88	1·08	0·35	0·66	0·87	0·83	0·85	0·69
50	1·21	1·34	0·79	1·18	1·26	1·17	1·24	1·05
$\Sigma(x)$	0·02	0·65	−5·90	−4·17	−1·18	−0·17	−1·06	0·12
$\Sigma(x^2)$	36·5678	44·7135	13·5620	26·8849	36·8026	34·5703	34·4902	28·4536

Eleventh fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	0·88	1·08	0·35	0·66	0·87	0·83	0·85	0·69
2	−0·62	−0·81	−0·51	−0·69	−0·69	−0·58	−0·66	−0·44
3	−0·36	−0·49	−0·44	−0·55	−0·44	−0·33	−0·44	−0·23
4	0·86	1·06	0·33	0·64	0·85	0·81	0·83	0·67
5	0·98	1·17	0·47	0·81	0·99	0·93	0·97	0·79
6	−0·73	−0·93	−0·54	−0·74	−0·79	−0·68	−0·76	−0·54
7	−0·52	−0·69	−0·49	−0·64	−0·59	−0·48	−0·58	−0·36
8	−0·72	−0·92	−0·54	−0·73	−0·78	−0·67	−0·75	−0·53
9	−0·97	−1·16	−0·59	−0·82	−1·00	−0·92	−0·95	−0·78
10	1·28	1·38	0·90	1·30	1·34	1·25	1·33	1·14
11	−0·18	−0·25	−0·38	−0·44	−0·26	−0·17	−0·27	−0·11
12	−0·27	−0·37	−0·41	−0·50	−0·35	−0·25	−0·36	−0·17
13	1·17	1·31	0·73	1·1[illegible]	1·21	1·13	1·20	1·00
14	0·91	1·10	0·39	0·71	0·91	0·86	0·89	0·72
15	0·15	0·21	−0·23	−0·18	0·08	0·14	0·05	0·09
16	0·72	0·92	0·19	0·44	0·69	0·67	0·67	0·53
17	−1·25	−1·37	−0·63	−0·89	−1·24	−1·21	−1·16	−1·10
18	−1·08	−1·25	−0·61	−0·85	−1·10	−1·04	−1·03	−0·90
19	−1·18	−1·32	−0·62	−0·87	−1·18	−1·14	−1·11	−1·02
20	−0·89	−1·09	−0·57	−0·79	−0·93	−0·84	−0·88	−0·70
21	0·30	0·41	−0·14	−0·04	0·24	0·28	0·21	0·19
22	−0·65	−0·84	−0·52	−0·70	−0·71	−0·61	−0·69	−0·47
23	0·88	1·08	0·35	0·66	0·87	0·83	0·85	0·69
24	0·98	1·17	0·47	0·81	0·99	0·93	0·97	0·79
25	1·97	1·65	2·56	2·71	2·19	2·03	2·20	2·14
26	−0·82	−1·02	−0·56	−0·77	−0·87	−0·77	−0·83	−0·63
27	0·60	0·78	0·08	0·29	0·56	0·56	0·53	0·42
28	−0·37	−0·50	−0·44	−0·56	−0·45	−0·34	−0·45	−0·24
29	0·68	0·87	0·15	0·39	0·65	0·64	0·62	0·49
30	−0·88	−1·08	−0·57	−0·79	−0·92	−0·83	−0·88	−0·69
31	1·73	1·59	1·85	2·17	1·89	1·75	1·89	1·75
32	0·38	0·51	−0·09	0·04	0·32	0·35	0·29	0·25
33	−0·65	−0·84	−0·52	−0·70	−0·71	−0·61	−0·69	−0·47
34	1·07	1·24	0·59	0·95	1·09	1·03	1·08	0·89
35	−0·05	−0·07	−0·32	−0·35	−0·13	−0·05	−0·15	−0·03
36	−0·54	−0·71	−0·49	−0·65	−0·61	−0·50	−0·60	−0·37
37	−0·28	−0·38	−0·41	−0·51	−0·36	−0·26	−0·36	−0·17
38	−0·61	−0·79	−0·51	−0·68	−0·68	−0·57	−0·66	−0·43
39	−0·22	−0·30	−0·39	−0·47	−0·30	−0·20	−0·31	−0·14
40	0·15	0·21	−0·23	−0·18	0·08	0·14	0·05	0·09
41	−0·86	−1·06	−0·57	−0·78	−0·91	−0·81	−0·86	−0·67
42	0·75	0·95	0·22	0·48	0·73	0·70	0·70	0·56
43	0·82	1·02	0·29	0·58	0·81	0·77	0·78	0·63
44	−0·73	−0·93	−0·54	−0·74	−0·79	−0·68	−0·76	−0·54
45	0·40	0·54	−0·07	0·07	0·34	0·37	0·32	0·26
46	0·17	0·23	−0·21	−0·16	0·10	0·16	0·07	0·10
47	1·36	1·43	1·04	1·44	1·44	1·33	1·43	1·24
48	−0·98	−1·17	−0·59	−0·82	−1·01	−0·93	−0·95	−0·79
49	−1·30	−1·40	−0·64	−0·90	−1·28	−1·27	−1·19	−1·16
50	0·06	0·08	−0·27	−0·26	−0·02	0·06	−0·04	0·03
$\Sigma(x)$	1·54	0·25	−3·68	−2·49	0·14	1·81	0·41	2·47
$\Sigma(x^2)$	35·9004	46·4311	21·2352	35·1287	39·0552	33·5157	37·0161	26·6425

Twelfth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	−1·01	−1·19	−0·59	−0·83	−1·04	−0·96	−0·98	−0·82
2	−0·43	−0·58	−0·46	−0·59	−0·50	−0·40	−0·50	−0·29
3	−0·76	−0·96	−0·55	−0·75	−0·82	−0·71	−0·78	−0·57
4	0·13	0·18	−0·24	−0·20	0·06	0·12	0·03	0·08
5	−0·87	−1·07	−0·57	−0·79	−0·92	−0·82	−0·87	−0·68
6	1·69	1·57	1·74	2·09	1·84	1·70	1·84	1·69
7	0·02	0·03	−0·29	−0·29	−0·06	0·02	−0·08	0·01
8	0·33	0·45	−0·12	−0·01	0·27	0·30	0·24	0·21
9	−0·79	−0·99	−0·55	−0·76	−0·84	−0·74	−0·80	−0·60
10	1·41	1·46	1·13	1·53	1·50	1·39	1·49	1·30
11	1·79	1·60	2·01	2·30	1·96	1·82	1·97	1·85
12	1·38	1·44	1·08	1·48	1·46	1·35	1·45	1·26
13	0·71	0·90	0·18	0·43	0·68	0·66	0·66	0·52
14	−0·28	−0·38	−0·41	−0·51	−0·36	−0·26	−0·36	−0·17
15	−0·96	−1·15	−0·59	−0·82	−0·99	−0·91	−0·94	−0·77
16	−0·09	−0·12	−0·34	−0·38	−0·17	−0·08	−0·19	−0·05
17	−0·60	−0·78	−0·51	−0·68	−0·67	−0·56	−0·65	−0·42
18	−0·09	−0·12	−0·34	−0·38	−0·17	−0·08	−0·19	−0·05
19	−0·51	−0·68	−0·49	−0·64	−0·58	−0·47	−0·57	−0·35
20	−0·29	−0·40	−0·42	−0·51	−0·37	−0·27	−0·37	−0·18
21	1·50	1·50	1·31	1·71	1·61	1·49	1·60	1·42
22	−0·47	−0·63	−0·47	−0·62	−0·54	−0·44	−0·54	−0·32
23	−2·34	−1·70	−0·72	−0·99	−1·84	−2·48	−1·99	−2·80
24	−1·09	−1·25	−0·61	−0·85	−1·11	−1·05	−1·04	−0·91
25	1·18	1·32	0·74	1·13	1·22	1·14	1·21	1·02
26	−0·53	−0·70	−0·49	−0·65	−0·60	−0·49	−0·59	−0·37
27	−0·73	−0·93	−0·54	−0·74	−0·79	−0·68	−0·76	−0·54
28	1·47	1·49	1·25	1·65	1·57	1·45	1·56	1·38
29	−0·78	−0·98	−0·55	−0·75	−0·83	−0·73	−0·80	−0·59
30	−0·30	−0·41	−0·42	−0·52	−0·38	−0·28	−0·38	−0·19
31	1·73	1·59	1·85	2·17	1·89	1·75	1·89	1·75
32	1·02	1·20	0·52	0·87	1·03	0·97	1·02	0·83
33	−0·31	−0·42	−0·42	−0·52	−0·39	−0·29	−0·39	−0·19
34	1·28	1·38	0·90	1·30	1·34	1·25	1·33	1·14
35	−1·14	−1·29	−0·61	−0·86	−1·15	−1·10	−1·08	−0·97
36	−0·27	−0·37	−0·41	−0·50	−0·35	−0·25	−0·36	−0·17
37	−1·14	−1·29	−0·61	−0·86	−1·15	−1·10	−1·08	−0·97
38	1·68	1·57	1·72	2·07	1·83	1·69	1·83	1·68
39	−1·70	−1·58	−0·68	−0·95	−1·56	−1·71	−1·48	−1·71
40	0·27	0·37	−0·16	−0·07	0·20	0·25	0·18	0·17
41	0·69	0·88	0·16	0·41	0·66	0·65	0·63	0·50
42	−0·85	−1·05	−0·57	−0·78	−0·90	−0·80	−0·85	−0·66
43	0·50	0·66	0·00	0·18	0·45	0·46	0·42	0·34
44	−0·05	−0·07	−0·32	−0·35	−0·13	−0·05	−0·15	−0·03
45	−0·15	−0·21	−0·36	−0·42	−0·23	−0·14	−0·24	−0·09
46	1·66	1·56	1·67	2·03	1·80	1·67	1·80	1·65
47	0·76	0·96	0·23	0·50	0·74	0·71	0·71	0·57
48	−1·38	−1·44	−0·65	−0·91	−1·34	−1·35	−1·25	−1·26
49	0·53	0·70	0·02	0·21	0·48	0·49	0·46	0·37
50	−0·87	−1·07	−0·57	−0·79	−0·92	−0·82	−0·87	−0·68
$\Sigma(x)$	0·95	−1·00	0·88	1·79	0·89	1·31	1·19	2·34
$\Sigma(x^2)$	51·6413	55·3842	31·7714	51·7893	54·4177	50·5009	52·9947	46·1012

Thirteenth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	0·30	0·41	−0·14	−0·04	0·24	0·28	0·21	0·19
2	0·95	1·14	0·43	0·77	0·95	0·90	0·93	0·76
3	−1·00	−1·18	−0·59	−0·83	−1·03	−0·95	−0·97	−0·81
4	1·43	1·47	1·17	1·57	1·52	1·41	1·51	1·33
5	1·06	1·23	0·57	0·93	1·08	1·01	1·06	0·88
6	−0·66	−0·85	−0·52	−0·71	−0·72	−0·62	−0·70	−0·48
7	0·60	0·78	0·08	0·29	0·56	0·56	0·53	0·42
8	1·11	1·27	0·64	1·01	1·14	1·07	1·12	0·93
9	−0·08	−0·11	−0·34	−0·37	−0·16	−0·07	−0·18	−0·05
10	1·41	1·46	1·13	1·53	1·50	1·39	1·49	1·30
11	0·78	0·98	0·25	0·52	0·76	0·73	0·74	0·59
12	−2·01	−1·66	−0·70	−0·98	−1·73	−2·08	−1·71	−2·20
13	−1·53	−1·51	−0·66	−0·93	−1·45	−1·52	−1·36	−1·46
14	0·30	0·41	−0·14	−0·04	0·24	0·28	0·21	0·19
15	0·87	1·07	0·34	0·65	0·86	0·82	0·84	0·68
16	−0·77	−0·97	−0·55	−0·75	−0·82	−0·72	−0·79	−0·58
17	2·32	1·70	3·94	3·59	2·63	2·46	2·68	2·75
18	0·62	0·81	0·10	0·32	0·58	0·58	0·56	0·44
19	−0·93	−1·12	−0·58	−0·81	−0·97	−0·88	−0·92	−0·74
20	1·65	1·56	1·65	2·01	1·79	1·66	1·79	1·64
21	0·13	0·18	−0·24	−0·20	0·06	0·12	0·03	0·08
22	1·35	1·43	1·02	1·42	1·42	1·32	1·41	1·22
23	−1·14	−1·29	−0·61	−0·86	−1·15	−1·10	−1·08	−0·97
24	0·35	0·47	−0·11	0·01	0·29	0·32	0·26	0·23
25	−2·15	−1·68	−0·71	−0·98	−1·79	−2·25	−1·83	−2·44
26	0·43	0·58	−0·05	0·10	0·37	0·40	0·35	0·29
27	−1·41	−1·46	−0·65	−0·92	−1·36	−1·39	−1·27	−1·30
28	−0·14	−0·19	−0·36	−0·41	−0·22	−0·13	−0·23	−0·08
29	1·69	1·57	1·74	2·09	1·84	1·70	1·84	1·69
30	0·84	1·04	0·31	0·61	0·83	0·79	0·81	0·65
31	−2·14	−1·68	−0·71	−0·98	−1·79	−2·24	−1·82	−2·43
32	−0·75	−0·95	−0·54	−0·74	−0·81	−0·70	−0·77	−0·56
33	−0·58	−0·76	−0·50	−0·67	−0·65	−0·54	−0·63	−0·41
34	−0·25	−0·34	−0·40	−0·49	−0·33	−0·23	−0·34	−0·16
35	0·29	0·40	−0·14	−0·05	0·23	0·27	0·20	0·18
36	−1·00	−1·18	−0·59	−0·83	−1·03	−0·95	−0·97	−0·81
37	−1·69	−1·57	−0·68	−0·95	−1·56	−1·70	−1·47	−1·69
38	0·63	0·82	0·11	0·33	0·59	0·59	0·57	0·45
39	−1·34	−1·42	−0·64	−0·91	−1·31	−1·31	−1·22	−1·21
40	−1·76	−1·60	−0·68	−0·96	−1·60	−1·78	−1·52	−1·80
41	0·73	0·93	0·20	0·46	0·70	0·68	0·68	0·54
42	−0·10	−0·14	−0·34	−0·38	−0·18	−0·09	−0·20	−0·06
43	0·69	0·88	0·16	0·41	0·66	0·65	0·63	0·50
44	0·63	0·82	0·11	0·33	0·59	0·59	0·57	0·45
45	−1·33	−1·41	−0·64	−0·90	−1·30	−1·30	−1·21	−1·20
46	−0·63	−0·82	−0·52	−0·69	−0·69	−0·59	−0·67	−0·45
47	0·11	0·15	−0·25	−0·22	0·04	0·10	0·01	0·06
48	−1·08	−1·25	−0·61	−0·85	−1·10	−1·04	−1·03	−0·90
49	1·05	1·22	0·56	0·92	1·07	1·00	1·05	0·87
50	0·48	0·64	−0·02	0·15	0·43	0·45	0·40	0·33
$\Sigma(x)$	−1·67	0·28	0·30	1·57	−0·78	−2·05	−0·41	−3·15
$\Sigma(x^2)$	61·4693	62·0998	34·5500	48·1979	59·7628	61·2141	57·3189	58·3595

Fourteenth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	0·84	1·04	0·31	0·61	0·83	0·79	0·81	0·65
2	0·77	0·97	0·24	0·51	0·75	0·72	0·73	0·58
3	0·78	0·98	0·25	0·52	0·76	0·73	0·74	0·59
4	1·78	1·60	1·98	2·28	1·95	1·81	1·96	1·83
5	−0·40	−0·54	−0·45	−0·58	−0·48	−0·37	−0·47	−0·26
6	0·08	0·11	−0·26	−0·24	0·01	0·07	−0·02	0·05
7	−0·01	−0·01	−0·30	−0·31	−0·09	−0·01	−0·11	−0·01
8	−0·14	−0·19	−0·36	−0·41	−0·22	−0·13	−0·23	−0·08
9	0·11	0·15	−0·25	−0·22	0·04	0·10	0·01	0·06
10	0·19	0·26	−0·20	−0·14	0·12	0·17	0·09	0·12
11	−0·09	−0·12	−0·34	−0·38	−0·17	−0·08	−0·19	−0·05
12	−0·40	−0·54	−0·45	−0·58	−0·48	−0·37	−0·47	−0·26
13	0·46	0·61	−0·03	0·13	0·41	0·43	0·38	0·31
14	0·07	0·10	−0·27	−0·25	0·00	0·06	−0·03	0·04
15	−1·95	−1·64	−0·70	−0·97	−1·70	−2·01	−1·67	−2·10
16	−1·09	−1·25	−0·61	−0·85	−1·11	−1·05	−1·04	−0·91
17	1·54	1·52	1·40	1·79	1·65	1·53	1·65	1·48
18	0·45	0·60	−0·04	0·12	0·40	0·42	0·37	0·30
19	−0·61	−0·79	−0·51	−0·68	−0·68	−0·57	−0·66	−0·43
20	−0·53	−0·70	−0·49	−0·65	−0·60	−0·49	−0·59	−0·37
21	0·54	0·71	0·03	0·22	0·49	0·50	0·47	0·37
22	1·49	1·50	1·29	1·69	1·59	1·48	1·59	1·41
23	0·53	0·70	0·02	0·21	0·48	0·49	0·46	0·37
24	2·40	1·70	4·34	3·80	2·73	2·55	2·79	2·91
25	−0·99	−1·17	−0·59	−0·82	−1·02	−0·94	−0·96	−0·80
26	0·79	0·99	0·26	0·54	0·77	0·74	0·75	0·60
27	0·07	0·10	−0·27	−0·25	0·00	0·06	−0·03	0·04
28	0·82	1·02	0·29	0·58	0·81	0·77	0·78	0·63
29	0·69	0·88	0·16	0·41	0·66	0·65	0·63	0·50
30	−0·44	−0·59	−0·46	−0·60	−0·51	−0·41	−0·51	−0·29
31	1·20	1·33	0·77	1·16	1·25	1·16	1·23	1·04
32	1·03	1·21	0·53	0·89	1·05	0·98	1·03	0·84
33	−1·25	−1·37	−0·63	−0·89	−1·24	−1·21	−1·16	−1·10
34	−0·89	−1·09	−0·57	−0·79	−0·93	−0·84	−0·88	−0·70
35	0·78	0·98	0·25	0·52	0·76	0·73	0·74	0·59
36	2·04	1·66	2·79	2·88	2·23	2·11	2·30	2·25
37	0·81	1·01	0·28	0·57	0·79	0·76	0·77	0·62
38	−0·93	−1·12	−0·58	−0·81	−0·97	−0·88	−0·92	−0·74
39	0·53	0·70	0·02	0·21	0·48	0·49	0·46	0·37
40	2·23	1·69	3·54	3·35	2·52	2·35	2·55	2·59
41	1·27	1·38	0·88	1·28	1·33	1·23	1·32	1·12
42	0·28	0·38	−0·15	−0·06	0·21	0·26	0·19	0·18
43	−1·38	−1·44	−0·65	−0·91	−1·34	−1·35	−1·25	−1·26
44	−0·77	−0·97	−0·55	−0·75	−0·82	−0·72	−0·79	−0·58
45	0·93	1·12	0·41	0·74	0·93	0·88	0·91	0·74
46	2·90	1·73	7·64	5·27	3·32	3·15	3·55	3·94
47	0·62	0·81	0·10	0·32	0·58	0·58	0·56	0·44
48	−1·30	−1·40	−0·64	−0·90	−1·28	−1·27	−1·19	−1·16
49	−0·08	−0·11	−0·34	−0·37	−0·16	−0·07	−0·18	−0·05
50	0·04	0·06	−0·28	−0·27	−0·04	0·04	−0·06	0·02
$\Sigma(x)$	15·81	14·56	16·81	16·92	16·11	16·02	16·41	16·43
$\Sigma(x^2)$	57·3619	52·9306	112·8177	89·2754	64·7851	58·6566	65·1299	62·5201

Fifteenth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	−0·01	−0·01	−0·30	−0·31	−0·09	−0·01	−0·11	−0·01
2	0·08	0·11	−0·26	−0·24	0·01	0·07	−0·02	0·05
3	2·51	1·71	4·93	4·12	2·86	2·69	2·95	3·13
4	−0·88	−1·08	−0·57	−0·79	−0·92	−0·83	−0·88	−0·69
5	0·64	0·83	0·11	0·34	0·60	0·60	0·58	0·46
6	0·14	0·19	−0·23	−0·19	0·07	0·13	0·04	0·08
7	0·82	1·02	0·29	0·58	0·81	0·77	0·78	0·63
8	−0·51	−0·68	−0·49	−0·64	−0·58	−0·47	−0·57	−0·35
9	−1·75	−1·59	−0·68	−0·96	−1·59	−1·77	−1·52	−1·78
10	0·47	0·63	−0·02	0·14	0·42	0·44	0·39	0·32
11	1·13	1·28	0·67	1·05	1·16	1·09	1·15	0·96
12	1·13	1·28	0·67	1·05	1·16	1·09	1·15	0·96
13	0·40	0·54	−0·07	0·07	0·34	0·37	0·32	0·26
14	−1·77	−1·60	−0·68	−0·96	−1·61	−1·79	−1·53	−1·81
15	−0·06	−0·08	−0·33	−0·35	−0·14	−0·06	−0·16	−0·03
16	0·47	0·63	−0·02	0·14	0·42	0·44	0·39	0·32
17	−0·18	−0·25	−0·38	−0·44	−0·26	−0·17	−0·27	−0·11
18	0·30	0·41	−0·14	−0·04	0·24	0·28	0·21	0·19
19	0·14	0·19	−0·23	−0·19	0·07	0·13	0·04	0·08
20	0·54	0·71	0·03	0·22	0·49	0·50	0·47	0·37
21	−0·07	−0·10	−0·33	−0·36	−0·15	−0·06	−0·17	−0·04
22	0·17	0·23	−0·21	−0·16	0·10	0·16	0·07	0·10
23	1·07	1·24	0·59	0·95	1·09	1·03	1·08	0·89
24	2·75	1·72	6·47	4·81	3·15	2·98	3·31	3·62
25	−2·67	−1·72	−0·73	−1·00	−1·86*	−2·88	−2·33	−3·45
26	−0·43	−0·58	−0·46	−0·59	−0·50	−0·40	−0·50	−0·29
27	−0·03	−0·04	−0·31	−0·33	−0·11	−0·03	−0·13	−0·02
28	−1·10	−1·26	−0·61	−0·85	−1·12	−1·06	−1·05	−0·92
29	−0·24	−0·33	−0·40	−0·48	−0·32	−0·22	−0·33	−0·15
30	0·47	0·63	−0·02	0·14	0·42	0·44	0·39	0·32
31	1·14	1·29	0·68	1·06	1·17	1·10	1·16	0·97
32	−0·64	−0·83	−0·52	−0·70	−0·70	−0·60	−0·68	−0·46
33	−0·46	−0·61	−0·47	−0·61	−0·53	−0·43	−0·53	−0·31
34	−0·12	−0·17	−0·35	−0·40	−0·20	−0·11	−0·21	−0·07
35	1·33	1·41	0·99	1·39	1·40	1·30	1·39	1·20
36	−1·05	−1·22	−0·60	−0·84	−1·07	−1·00	−1·01	−0·87
37	−1·20	−1·33	−0·62	−0·88	−1·20	−1·16	−1·12	−1·04
38	0·75	0·95	0·22	0·48	0·73	0·70	0·70	0·56
39	0·11	0·15	−0·25	−0·22	0·04	0·10	0·01	0·06
40	0·44	0·59	−0·04	0·11	0·39	0·41	0·36	0·29
41	−2·39	−1·70	−0·72	−0·99	−1·85	−2·54	−2·04	−2·89
42	0·95	1·14	0·43	0·77	0·95	0·90	0·93	0·76
43	−0·54	−0·71	−0·49	−0·65	−0·61	−0·50	−0·60	−0·37
44	−0·93	−1·12	−0·58	−0·81	−0·97	−0·88	−0·92	−0·74
45	−0·86	−1·06	−0·57	−0·78	−0·91	−0·81	−0·86	−0·67
46	−0·97	−1·16	−0·59	−0·82	−1·00	−0·92	−0·95	−0·78
47	−1·28	−1·38	−0·63	−0·89	−1·26	−1·25	−1·18	−1·14
48	0·08	0·11	−0·26	−0·24	0·01	0·07	−0·02	0·05
49	−0·12	−0·17	−0·35	−0·40	−0·20	−0·11	−0·21	−0·07
50	−0·50	−0·66	−0·48	−0·63	−0·57	−0·46	−0·56	−0·34
$\Sigma(x)$	−2·73	−2·45	1·09	−1·32	−2·22	−2·73	−2·59	−2·77
$\Sigma(x^2)$	54·1509	46·5775	76·7501	60·9554	52·2892	56·5775	54·6291	63·6593

Sixteenth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	0·24	0·33	−0·17	−0·10	0·17	0·22	0·15	0·15
2	−0·57	−0·75	−0·50	−0·67	−0·64	−0·53	−0·62	−0·40
3	−0·72	−0·92	−0·54	−0·73	−0·78	−0·67	−0·75	−0·53
4	−0·30	−0·41	−0·42	−0·52	−0·38	−0·28	−0·38	−0·19
5	0·88	1·08	0·35	0·66	0·87	0·83	0·85	0·69
6	0·30	0·41	−0·14	−0·04	0·24	0·28	0·21	0·19
7	1·28	1·38	0·90	1·30	1·34	1·25	1·33	1·14
8	−1·38	−1·44	−0·65	−0·91	−1·34	−1·35	−1·25	−1·26
9	1·68	1·57	1·72	2·07	1·83	1·69	1·83	1·68
10	−0·17	−0·23	−0·37	−0·43	−0·25	−0·16	−0·26	−0·10
11	−0·14	−0·19	−0·36	−0·41	−0·22	−0·13	−0·23	−0·08
12	0·68	0·87	0·15	0·39	0·65	0·64	0·62	0·50
13	0·18	0·25	−0·21	−0·15	0·11	0·17	0·08	0·11
14	−1·71	−1·58	−0·68	−0·96	−1·57	−1·72	−1·49	−1·73
15	−0·28	−0·38	−0·41	−0·51	−0·36	−0·26	−0·36	−0·18
16	0·94	1·13	0·42	0·75	0·94	0·89	0·92	0·75
17	0·33	0·45	−0·12	−0·01	0·27	0·30	0·24	0·21
18	1·18	1·32	0·74	1·13	1·22	1·14	1·21	1·02
19	−0·62	−0·81	−0·51	−0·69	−0·69	−0·58	−0·66	−0·44
20	1·69	1·57	1·74	2·09	1·84	1·70	1·84	1·69
21	−0·40	−0·54	−0·45	−0·58	−0·48	−0·37	−0·47	−0·26
22	−1·96	−1·65	−0·70	−0·97	−1·71	−2·02	−1·67	−2·12
23	−0·84	−1·04	−0·56	−0·78	−0·89	−0·79	−0·84	−0·65
24	−0·54	−0·71	−0·49	−0·65	−0·61	−0·50	−0·60	−0·37
25	0·05	0·07	−0·28	−0·27	−0·03	0·05	−0·05	0·03
26	−0·82	−1·02	−0·56	−0·77	−0·87	−0·77	−0·83	−0·63
27	0·46	0·61	−0·03	0·13	0·41	0·43	0·38	0·31
28	−0·41	−0·55	−0·46	−0·58	−0·49	−0·38	−0·48	−0·27
29	1·27	1·38	0·88	1·28	1·33	1·23	1·32	1·12
30	−1·22	−1·35	−0·63	−0·88	−1·21	−1·18	−1·13	−1·06
31	1·76	1·60	1·93	2·24	1·93	1·78	1·93	1·80
32	−0·20	−0·27	−0·38	−0·45	−0·28	−0·18	−0·29	−0·12
33	1·05	1·22	0·56	0·92	1·07	1·00	1·05	0·87
34	−0·95	−1·14	−0·58	−0·81	−0·99	−0·90	−0·93	−0·76
35	−0·16	−0·22	−0·37	−0·43	−0·24	−0·15	−0·25	−0·10
36	−1·11	−1·27	−0·61	−0·86	−1·12	−1·07	−1·05	−0·93
37	0·97	1·16	0·46	0·80	0·98	0·92	0·96	0·78
38	1·53	1·51	1·37	1·76	1·64	1·52	1·64	1·46
39	−0·05	−0·07	−0·32	−0·35	−0·13	−0·05	−0·15	−0·03
40	0·64	0·83	0·11	0·34	0·60	0·60	0·58	0·46
41	−1·83	−1·62	−0·69	−0·97	−1·64	−1·86	−1·58	−1·91
42	−0·73	−0·93	−0·54	−0·74	−0·79	−0·68	−0·76	−0·54
43	1·26	1·37	0·87	1·27	1·32	1·22	1·30	1·11
44	−0·11	−0·15	−0·35	−0·39	−0·19	−0·10	−0·20	−0·06
45	0·02	0·03	−0·29	−0·29	−0·06	0·02	−0·08	0·01
46	−1·52	−1·51	−0·66	−0·93	−1·44	−1·51	−1·35	−1·45
47	−1·61	−1·55	−0·67	−0·94	−1·51	−1·61	−1·42	−1·58
48	−0·83	−1·03	−0·56	−0·77	−0·88	−0·78	−0·84	−0·64
49	1·78	1·60	1·98	2·28	1·95	1·81	1·96	1·83
50	0·29	0·40	−0·14	−0·05	0·23	0·27	0·20	0·18
$\Sigma(x)$	−0·72	−1·19	−1·22	−0·18	−0·85	−0·62	−0·37	−0·30
$\Sigma(x^2)$	50·9664	54·6567	27·1922	45·1382	52·4567	49·6494	49·9199	44·7696

Seventeenth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	−0·75	−0·95	−0·54	−0·74	−0·81	−0·70	−0·77	−0·56
2	−0·47	−0·63	−0·47	−0·62	−0·54	−0·44	−0·54	−0·32
3	0·55	0·72	0·04	0·23	0·51	0·51	0·48	0·38
4	0·87	1·07	0·34	0·65	0·86	0·82	0·84	0·68
5	−0·63	−0·82	−0·52	−0·69	−0·69	−0·59	−0·67	−0·45
6	−0·60	−0·78	−0·51	−0·68	−0·67	−0·56	−0·65	−0·42
7	0·85	1·05	0·32	0·62	0·84	0·80	0·82	0·66
8	−0·86	−1·06	−0·57	−0·78	−0·91	−0·81	−0·86	−0·67
9	0·79	0·99	0·26	0·54	0·77	0·74	0·75	0·60
10	0·77	0·97	0·24	0·51	0·75	0·72	0·73	0·58
11	−1·31	−1·40	−0·64	−0·90	−1·29	−1·28	−1·20	−1·17
12	1·02	1·20	0·52	0·87	1·03	0·97	1·02	0·83
13	1·01	1·19	0·51	0·86	1·02	0·96	1·00	0·82
14	−0·69	−0·88	−0·53	−0·72	−0·75	−0·65	−0·72	−0·50
15	−0·59	−0·77	−0·51	−0·67	−0·66	−0·55	−0·64	−0·42
16	−1·76	−1·60	−0·68	−0·96	−1·60	−1·78	−1·52	−1·80
17	0·47	0·63	−0·02	0·14	0·42	0·44	0·39	0·32
18	2·36	1·70	4·14	3·70	2·68	2·50	2·74	2·83
19	−0·88	−1·08	−0·57	−0·79	−0·92	−0·83	−0·88	−0·69
20	−0·42	−0·56	−0·46	−0·59	−0·50	−0·39	−0·49	−0·28
21	0·06	0·08	−0·27	−0·26	−0·02	0·06	−0·04	0·03
22	−1·55	−1·52	−0·66	−0·94	−1·46	−1·54	−1·37	−1·49
23	0·92	1·11	0·40	0·72	0·92	0·87	0·90	0·73
24	0·48	0·64	−0·02	0·15	0·43	0·45	0·40	0·33
25	0·88	1·08	0·35	0·66	0·87	0·83	0·85	0·69
26	−2·01	−1·66	−0·70	−0·98	−1·73	−2·08	−1·71	−2·20
27	1·62	1·55	1·57	1·95	1·75	1·62	1·75	1·59
28	0·40	0·54	−0·07	0·07	0·34	0·37	0·32	0·26
29	1·94	1·64	2·46	2·64	2·15	1·99	2·16	2·09
30	0·49	0·65	−0·01	0·16	0·44	0·45	0·41	0·33
31	0·36	0·49	−0·10	0·02	0·30	0·33	0·27	0·23
32	−0·47	−0·63	−0·47	−0·62	−0·54	−0·44	−0·54	−0·32
33	2·19	1·68	3·37	3·25	2·47	2·30	2·50	2·51
34	1·22	1·35	0·80	1·20	1·27	1·18	1·26	1·06
35	0·69	0·88	0·16	0·41	0·66	0·65	0·63	0·50
36	−0·70	−0·89	−0·53	−0·72	−0·76	−0·66	−0·73	−0·51
37	−1·27	−1·38	−0·63	−0·89	−1·25	−1·23	−1·17	−1·12
38	0·18	0·25	−0·21	−0·15	0·11	0·17	0·08	0·11
39	0·98	1·17	0·47	0·81	0·99	0·93	0·97	0·79
40	−0·76	−0·96	−0·55	−0·75	−0·82	−0·71	−0·78	−0·57
41	−0·48	−0·64	−0·48	−0·62	−0·55	−0·45	−0·54	−0·33
42	1·49	1·50	1·29	1·69	1·59	1·48	1·59	1·41
43	−0·34	−0·46	−0·43	−0·54	−0·42	−0·31	−0·42	−0·22
44	0·15	0·21	−0·23	−0·18	0·08	0·14	0·05	0·09
45	−0·77	−0·97	−0·55	−0·75	−0·82	−0·72	−0·79	−0·58
46	1·28	1·38	0·90	1·30	1·34	1·25	1·33	1·14
47	−1·16	−1·31	−0·62	−0·87	−1·16	−1·12	−1·09	−0·99
48	1·88	1·63	2·27	2·50	2·08	1·92	2·09	1·99
49	−0·67	−0·86	−0·53	−0·71	−0·73	−0·63	−0·71	−0·49
50	−0·78	−0·98	−0·55	−0·75	−0·83	−0·73	−0·80	−0·59
$\Sigma(x)$	5·98	4·56	6·78	7·78	6·24	6·25	6·70	6·89
$\Sigma(x^2)$	56·3456	58·5270	53·9998	65·4558	60·9584	55·8065	59·2344	52·9191

Eighteenth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	−0·97	−1·16	−0·59	−0·82	−1·00	−0·92	−0·95	−0·78
2	2·38	1·70	4·24	3·74	2·70	2·53	2·76	2·86
3	0·78	0·98	0·25	0·52	0·76	0·73	0·74	0·59
4	0·35	0·47	−0·11	0·01	0·29	0·32	0·26	0·23
5	1·10	1·26	0·63	1·00	1·13	1·06	1·11	0·92
6	−0·03	−0·04	−0·31	−0·33	−0·11	−0·03	−0·13	−0·02
7	−1·47	−1·49	−0·66	−0·93	−1·41	−1·45	−1·32	−1·38
8	0·68	0·87	0·15	0·39	0·65	0·64	0·62	0·50
9	−0·73	−0·93	−0·54	−0·74	−0·79	−0·68	−0·76	−0·54
10	−1·17	−1·31	−0·62	−0·87	−1·17	−1·13	−1·10	−1·00
11	−0·05	−0·07	−0·32	−0·35	−0·13	−0·05	−0·15	−0·03
12	1·02	1·20	0·52	0·87	1·03	0·97	1·02	0·83
13	−0·19	−0·26	−0·38	−0·45	−0·27	−0·17	−0·28	−0·12
14	−0·89	−1·09	−0·57	−0·79	−0·93	−0·84	−0·88	−0·70
15	0·56	0·74	0·05	0·25	0·52	0·52	0·49	0·39
16	0·47	0·63	−0·02	0·14	0·42	0·44	0·39	0·32
17	−0·70	−0·89	−0·53	−0·72	−0·76	−0·66	−0·73	−0·51
18	1·49	1·50	1·29	1·69	1·59	1·48	1·59	1·41
19	−0·75	−0·95	−0·54	−0·74	−0·81	−0·70	−0·77	−0·56
20	1·03	1·21	0·53	0·89	1·05	0·98	1·03	0·84
21	0·24	0·33	−0·17	−0·10	0·17	0·22	0·15	0·15
22	−0·42	−0·56	−0·46	−0·59	−0·50	−0·39	−0·49	−0·28
23	−0·67	−0·86	−0·53	−0·71	−0·73	−0·63	−0·71	−0·49
24	−1·32	−1·41	−0·64	−0·90	−1·29	−1·29	−1·21	−1·19
25	0·20	0·27	−0·20	−0·13	0·13	0·18	0·10	0·12
26	1·31	1·40	0·95	1·35	1·38	1·28	1·37	1·17
27	1·11	1·27	0·64	1·01	1·14	1·07	1·12	0·93
28	−0·21	−0·29	−0·39	−0·46	−0·29	−0·19	−0·30	−0·13
29	−0·35	−0·47	−0·44	−0·55	−0·43	−0·32	−0·43	−0·23
30	−2·76	−1·72	−0·73	−1·00	−1·86*	−2·99	−2·44	−3·64
31	−0·63	−0·82	−0·52	−0·69	−0·69	−0·59	−0·67	−0·45
32	−2·31	−1·70	−0·72	−0·99	−1·83	−2·44	−1·96	−2·74
33	−0·01	−0·01	−0·30	−0·31	−0·09	−0·01	−0·11	−0·01
34	0·41	0·55	−0·07	0·08	0·35	0·38	0·33	0·27
35	3·02	1·73	8·72	5·65	3·45	3·29	3·74	4·21
36	1·26	1·37	0·87	1·27	1·32	1·22	1·30	1·11
37	0·36	0·49	−0·10	0·02	0·30	0·33	0·27	0·23
38	−0·43	−0·58	−0·46	−0·59	−0·50	−0·40	−0·50	−0·29
39	−0·26	−0·36	−0·41	−0·49	−0·34	−0·24	−0·35	−0·16
40	1·73	1·59	1·85	2·17	1·89	1·75	1·89	1·75
41	−0·64	−0·83	−0·52	−0·70	−0·70	−0·60	−0·68	−0·46
42	0·20	0·27	−0·20	−0·13	0·13	0·18	0·10	0·12
43	−2·08	−1·67	−0·71	−0·98	−1·76	−2·16	−1·77	−2·32
44	−0·46	−0·61	−0·47	−0·61	−0·53	−0·43	−0·53	−0·31
45	−0·84	−1·04	−0·56	−0·78	−0·89	−0·79	−0·84	−0·65
46	0·16	0·22	−0·22	−0·17	0·09	0·15	0·06	0·10
47	−1·42	−1·46	−0·65	−0·92	−1·37	−1·40	−1·28	−1·32
48	0·20	0·27	−0·20	−0·13	0·13	0·18	0·10	0·12
49	1·34	1·42	1·00	1·41	1·41	1·31	1·40	1·21
50	−0·63	−0·82	−0·52	−0·69	−0·69	−0·59	−0·67	−0·45
$\Sigma(x)$	−0·99	−1·66	6·31	3·10	0·16	−0·88	−0·07	−0·38
$\Sigma(x^2)$	62·9609	53·5924	111·1731	77·1048	62·0682	65·7264	65·6813	75·0474

Nineteenth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	−0·19	−0·26	−0·38	−0·45	−0·27	−0·17	−0·28	−0·12
2	0·13	0·18	−0·24	−0·20	0·06	0·12	0·03	0·08
3	−0·82	−1·02	−0·56	−0·77	−0·87	−0·77	−0·83	−0·63
4	−0·18	−0·25	−0·38	−0·44	−0·26	−0·17	−0·27	−0·11
5	−2·43	−1·71	−0·72	−0·99	−1·85	−2·59	−2·08	−2·97
6	0·78	0·98	0·25	0·52	0·76	0·73	0·74	0·59
7	−0·98	−1·17	−0·59	−0·82	−1·01	−0·93	−0·95	−0·79
8	−0·09	−0·12	−0·34	−0·38	−0·17	−0·08	−0·19	−0·05
9	−0·81	−1·01	−0·56	−0·77	−0·86	−0·76	−0·82	−0·62
10	0·56	0·74	0·05	0·25	0·52	0·52	0·49	0·39
11	−2·32	−1·70	−0·72	−0·99	−1·84	−2·46	−1·97	−2·75
12	0·01	0·01	−0·30	−0·30	−0·07	0·01	−0·09	0·01
13	−0·90	−1·09	−0·57	−0·80	−0·94	−0·85	−0·89	−0·71
14	0·22	0·30	−0·19	−0·12	0·15	0·20	0·13	0·14
15	−1·36	−1·43	−0·64	−0·91	−1·32	−1·33	−1·24	−1·24
16	0·65	0·84	0·12	0·36	0·62	0·61	0·59	0·47
17	0·39	0·53	−0·08	0·05	0·33	0·36	0·30	0·26
18	0·53	0·70	0·02	0·21	0·48	0·49	0·46	0·37
19	0·81	1·01	0·28	0·57	0·79	0·76	0·77	0·62
20	0·91	1·10	0·39	0·71	0·91	0·86	0·89	0·72
21	−0·14	−0·19	−0·36	−0·41	−0·22	−0·13	−0·23	−0·08
22	1·76	1·60	1·93	2·24	1·93	1·78	1·93	1·80
23	1·19	1·33	0·76	1·15	1·23	1·15	1·22	1·03
24	1·41	1·46	1·13	1·53	1·50	1·39	1·49	1·30
25	−1·00	−1·18	−0·59	−0·83	−1·03	−0·95	−0·97	−0·81
26	−1·20	−1·33	−0·62	−0·88	−1·20	−1·16	−1·12	−1·04
27	0·71	0·90	0·18	0·43	0·68	0·66	0·66	0·52
28	−0·59	−0·77	−0·51	−0·67	−0·66	−0·55	−0·64	−0·42
29	−1·38	−1·44	−0·65	−0·91	−1·34	−1·35	−1·25	−1·26
30	2·17	1·68	3·29	3·20	2·44	2·27	2·47	2·48
31	0·30	0·41	−0·14	−0·04	0·24	0·28	0·21	0·19
32	0·51	0·68	0·01	0·19	0·46	0·47	0·44	0·35
33	−0·07	−0·10	−0·33	−0·36	−0·15	−0·06	−0·17	−0·04
34	0·55	0·72	0·04	0·23	0·51	0·51	0·48	0·38
35	0·52	0·69	0·02	0·20	0·47	0·48	0·45	0·36
36	0·01	0·01	−0·30	−0·30	−0·07	0·01	−0·09	0·01
37	1·37	1·44	1·06	1·46	1·45	1·34	1·44	1·25
38	0·82	1·02	0·29	0·58	0·81	0·77	0·78	0·63
39	−0·08	−0·11	−0·34	−0·37	−0·16	−0·07	−0·18	−0·05
40	−0·25	−0·34	−0·40	−0·49	−0·33	−0·23	−0·34	−0·16
41	1·32	1·41	0·97	1·37	1·39	1·29	1·38	1·19
42	−0·50	−0·66	−0·48	−0·63	−0·57	−0·46	−0·56	−0·34
43	−0·04	−0·06	−0·32	−0·34	−0·12	−0·04	−0·14	−0·02
44	0·98	1·17	0·47	0·81	0·99	0·93	0·97	0·79
45	−0·50	−0·66	−0·48	−0·63	−0·57	0·46	−0·56	−0·34
46	0·51	0·68	0·01	0·19	0·46	0·47	0·44	0·35
47	1·59	1·54	1·51	1·88	1·72	1·59	1·71	1·55
48	0·44	0·59	−0·04	0·11	0·39	0·41	0·36	0·29
49	0·97	1·16	0·46	0·80	0·98	0·92	0·96	0·78
50	−1·29	−1·39	−0·64	−0·90	−1·27	−1·26	−1·19	−1·15
$\Sigma(x)$	5·00	6·89	0·77	3·34	5·12	5·47	4·74	3·20
$\Sigma(x^2)$	48·4038	49·4833	27·9639	40·9940	47·0018	48·2137	46·6414	46·8674

Twentieth fifty

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	1·31	1·40	0·95	1·35	1·38	1·28	1·37	1·17
2	−0·17	−0·23	−0·37	−0·43	−0·25	−0·16	−0·26	−0·10
3	−0·95	−1·14	−0·58	−0·81	−0·99	−0·90	−0·93	−0·76
4	0·69	0·88	0·16	0·41	0·66	0·65	0·63	0·50
5	0·32	0·43	−0·13	−0·02	0·26	0·30	0·23	0·20
6	0·26	0·36	−0·16	−0·08	0·19	0·24	0·17	0·16
7	−0·39	−0·53	−0·45	−0·57	−0·47	−0·36	−0·46	−0·26
8	−0·50	−0·66	−0·48	−0·63	−0·57	−0·46	−0·56	−0·34
9	1·91	1·63	2·36	2·57	2·11	1·96	2·13	2·04
10	2·37	1·70	4·19	3·72	2·69	2·52	2·75	2·85
11	1·77	1·60	1·95	2·26	1·94	1·79	1·94	1·81
12	−0·09	−0·12	−0·34	−0·38	−0·17	−0·08	−0·19	−0·05
13	−0·62	−0·81	−0·51	−0·69	−0·69	−0·58	−0·66	−0·44
14	−1·02	−1·20	−0·60	−0·83	−1·05	−0·97	−0·99	−0·83
15	−0·54	−0·71	−0·49	−0·65	−0·61	−0·50	−0·60	−0·37
16	0·07	0·10	−0·27	−0·25	0·00	0·06	−0·03	0·04
17	−0·16	−0·22	−0·37	−0·43	−0·24	−0·15	−0·25	−0·10
18	−1·44	−1·47	−0·65	−0·92	−1·39	−1·42	−1·29	−1·34
19	−0·60	−0·78	−0·51	−0·68	−0·67	−0·56	−0·65	−0·42
20	−0·23	−0·32	−0·40	−0·47	−0·31	−0·21	−0·32	−0·14
21	−0·67	−0·86	−0·53	−0·71	−0·73	−0·63	−0·71	−0·49
22	1·56	1·53	1·44	1·82	1·68	1·55	1·68	1·51
23	0·51	0·68	0·01	0·19	0·46	0·47	0·44	0·35
24	0·33	0·45	−0·12	−0·01	0·27	0·30	0·24	0·21
25	−0·01	−0·01	−0·30	−0·31	−0·09	−0·01	−0·11	−0·01
26	0·13	0·18	−0·24	−0·20	0·06	0·12	0·03	0·08
27	0·81	1·01	0·28	0·57	0·79	0·76	0·77	0·62
28	−1·09	−1·25	−0·61	−0·85	−1·11	−1·05	−1·04	−0·91
29	−0·90	−1·09	−0·57	−0·80	−0·94	−0·85	−0·89	−0·71
30	−0·39	−0·53	−0·45	−0·57	−0·47	−0·36	−0·46	−0·26
31	0·57	0·75	0·06	0·26	0·53	0·53	0·50	0·40
32	−0·88	−1·08	−0·57	−0·79	−0·92	−0·83	−0·88	−0·69
33	0·57	0·75	0·06	0·26	0·53	0·53	0·50	0·40
34	0·98	1·17	0·47	0·81	0·99	0·93	0·97	0·79
35	−0·38	−0·51	−0·45	−0·57	−0·46	−0·35	−0·46	−0·25
36	−1·06	−1·23	−0·60	−0·84	−1·08	−1·01	−1·02	−0·88
37	1·79	1·60	2·01	2·30	1·96	1·82	1·97	1·85
38	0·00	0·00	−0·30	−0·31	−0·08	0·00	−0·10	0·00
39	0·23	0·32	−0·18	−0·11	0·16	0·21	0·14	0·14
40	0·21	0·29	−0·19	−0·12	0·14	0·19	0·11	0·13
41	−0·76	−0·96	−0·55	−0·75	−0·82	−0·71	−0·78	−0·57
42	−0·94	−1·13	−0·58	−0·81	−0·98	−0·89	−0·92	−0·75
43	−0·98	−1·17	−0·59	−0·82	−1·01	−0·93	−0·95	−0·79
44	0·97	1·16	0·46	0·80	0·98	0·92	0·96	0·78
45	−0·64	−0·83	−0·52	−0·70	−0·70	−0·60	−0·68	−0·46
46	0·16	0·22	−0·22	−0·17	0·09	0·15	0·06	0·10
47	−1·39	−1·45	−0·65	−0·91	−1·35	−1·37	−1·26	−1·28
48	−0·03	−0·04	−0·31	−0·33	−0·11	−0·03	−0·13	−0·02
49	0·36	0·49	−0·10	0·02	0·30	0·33	0·27	0·23
50	−0·87	−1·07	−0·57	−0·79	−0·92	−0·82	−0·87	−0·68
$\Sigma(x)$	0·18	−2·70	−1·11	−1·97	−1·01	0·82	−0·59	2·46
$\Sigma(x^2)$	39·9808	43·9182	42·0127	51·0911	45·1523	39·0722	43·5785	35·7298

Note on Mr Quenouille's Edgeworth Type A transformation

By E. S. PEARSON

University College London

In representing a non-normal frequency function, $f(x)$, by a limited number of terms of an Edgeworth Type A expansion, statisticians have frequently used the form

$$f(x) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}x^2} \left\{1 + \frac{\kappa_3}{6}(x^3 - 3x) + \frac{\kappa_4}{24}(x^4 - 6x^2 + 3) + \frac{\kappa_3^2}{72}(x^6 - 15x^4 + 45x^2 - 15)\right\}. \quad (1)$$

This distribution, for example, formed the basis of Gayen's (1949, 1950) papers exploring the effect of departure from normality on several common statistical tests. It is also used by Srivastava in a paper printed elsewhere in this issue of *Biometrika*. It is important to realize, however, that the distributions of x_5, x_6 and x_7 derived from a normally distributed x_1 through the Cornish–Fisher equations given by Quenouille on p. 179 above, will not correspond exactly to particular cases of the frequency function (1). It seemed of interest to examine the extent of this difference.

(i) For the case $\kappa_3 = 0{\cdot}5$, $\kappa_4 = 0{\cdot}0$, $f(x)$ of equation (1) becomes negative* at $x_5 = -2{\cdot}53$, the negative area under the curve beyond this point amounting to $-0{\cdot}0028$. This means that the area above $-2{\cdot}53$ is 1·0028 or is in excess by 3 in 1000. For the Quenouille transformation the critical point is at $x_5 = -1{\cdot}86$. Thus there is inevitably some disagreement between the two distributions at the lower tail. It was found that at the upper tail there was good agreement† up to $x_5 = 3{\cdot}2$, beyond which x_5 for distribution (1) begins to increase steadily over Quenouille's value given in his table on p. 182.

(ii) For the case $\kappa_3 = 0{\cdot}0$, $\kappa_4 = 0{\cdot}5$, $f(x)$ for (1) does not become negative and the two x_6 distributions scarcely differ until after $x_6 = \pm 3{\cdot}4$.

(iii) For the case of $\kappa_3 = 0{\cdot}5$, $\kappa_4 = 0{\cdot}5$, the $f(x)$ of equation (1) becomes negative at $x_7 = -2{\cdot}65$ and positive again at about $x_7 = -3{\cdot}35$, but the effect of this anomaly is much smaller than for x_5. Since these extreme negative and positive areas almost cancel out, the area under the curve above $-2{\cdot}65$ is 0·99996 or only negligibly in error. In the lower tail, the values of x_7 obtained from (1) do not differ from those in Quenouille's table by more than 0·05 until $x_7 = -2{\cdot}1$; beyond this point Quenouille's x_7, for given x_1, assumes increasingly greater negative values than the x_7 of (1). However, as the table below shows, this difference is unlikely to be noticeable in a sample of 1000.

In this table are given the expected frequencies for the two asymmetrical Edgeworth distributions of equation (1), expressed in terms of totals of 1000. These are compared with Quenouille's two samples tabled on p. 181. The only appreciable difference is at the lower end of the distribution of x_5. For (1) there is an expected frequency of 8·6 in the interval $-2{\cdot}0$ to $-2{\cdot}53$, while no observations are possible in this range using Quenouille's transformation. However, if the χ^2 test for goodness of fit is used to compare the sampled frequencies with the expectations tabled below, pooling tail groups as shown by the braces, it is found that

(*a*) for $\kappa_3 = 0{\cdot}5$, $\kappa_4 = 0{\cdot}0$: $\chi^2 = 6{\cdot}29$, $\nu = 13$;

(*b*) for $\kappa_3 = 0{\cdot}5$, $\kappa_4 = 0{\cdot}5$: $\chi^2 = 15{\cdot}65$, $\nu = 14$.

Neither of these values is significant. One must conclude therefore that except at the lower tail of the x_5 distribution it would be difficult to detect any difference between Quenouille's three Edgeworth distributions and those derived from the commonly used form of equation (1), even in samples as large as 1000.

REFERENCES

Barton, D. E. & Dennis, K. E. R. (1952). *Biometrika*, **39**, 425–7.

Gayen, A. K. (1949). *Biometrika*, **36**, 353–69.

Gayen, A. K. (1950). *Biometrika*, **37**, 236–55, 399–408.

* This is to be expected from Barton & Dennis's (1952) investigation. $\beta_1 = \kappa_3^2$ and $\beta_2 = \kappa_4 + 3$ give a (β_1, β_2) point outside the region in which the $f(x)$ of (1) is positive definite.

† That is, no difference between the corresponding values of over 0·05.

Comparison of Edgeworth Type A distributions

x	Case $\kappa_3 = 0{\cdot}5$, $\kappa_4 = 0{\cdot}0$, i.e. distribution of x_5		Case $\kappa_3 = 0{\cdot}5$, $\kappa_4 = 0{\cdot}5$, i.e. distribution of x_7	
	Expectation from (1)	Quenouille's sample	Expectation from (1)	Quenouille's sample
$> 4{\cdot}0$	0·4	—	0·5	—
3·6 to 4·0	0·8	—	1·1	1
3·2 to 3·6	1·8	3	2·5	3
2·8 to 3·2	3·9	2	5·0	1
2·4 to 2·8	8·4	11	9·3	13
2·0 to 2·4	17·5	10	16·7	8
1·6 to 2·0	33·8	33	30·0	33
1·2 to 1·6	57·5	60	51·7	58
0·8 to 1·2	85·8	87	82·3	79
0·4 to 0·8	114·9	120	117·9	116
0·0 to 0·4	140·7	145	149·8	146
$-0{\cdot}4$ to 0·0	155·6	154	164·7	167
$-0{\cdot}8$ to $-0{\cdot}4$	149·0	154	152·1	174
$-1{\cdot}2$ to $-0{\cdot}8$	117·4	117	113·9	112
$-1{\cdot}6$ to $-1{\cdot}2$	72·0	66	66·3	56
$-2{\cdot}0$ to $-1{\cdot}6$	31·9	38	28·1	26
$-2{\cdot}4$ to $-2{\cdot}0$	8·3	—	7·5	5
$< -2{\cdot}4$*	0·3	—	0·6	2
Total	1000·0	1000	1000·0	1000

* For equation (1), the lower limit for x_5 is $-2{\cdot}53$ and for x_7 it is $-2{\cdot}65$.

MONOMIAL-MONOMIAL SYMMETRIC FUNCTION TABLES

By JACK LEVINE

University of North Carolina, Raleigh

1. The present monomial-monomial (*MM*) Table is a continuation of the David–Kendall Tables (1949), (1951), (1953), (1955). All monomial products up to and including weight $w = 8$ are listed as linear functions of monomials. Thus, Table 8, section (iii) and (iv), gives

$$(21)^2(2) = (62) + 2(61^2) + 2(53) + 2(521) + 2(4^2) + 4(431) + 8(422) + 6(421^2) + 6(3^22) + 8(32^21) + 24(2^4) + 12(2^31^2).$$

2. The use of *MM* expansions in problems of distributions is illustrated by a theorem of MacMahon.

The number of distributions of n objects of specification $[m_1 m_2 \dots m_l]$ in r different boxes $B_1, \dots, B_r$, such that B_i contains p_i objects of specification $[m_{1i} m_{2i} \dots m_{ki}]$ is given by the coefficient $C_{m_1 \dots m_l}$ of the monomial symmetric function $(m_1 \dots m_l)$ in the expansion of the product

$$(m_{11} m_{21} \dots m_{k_1 1})(m_{12} m_{22} \dots m_{k_2 2}) \dots (m_{1r} m_{2r} \dots m_{k_r r}) = \Sigma C_{q_1 \dots q_c}(q_1 \dots q_c),$$

the summation being over all partitions $[q_1 \dots q_c]$ of $n = \Sigma q_i = \Sigma p_i$.

3. Construction of the Tables was based on MacMahon's D-operator, the Tables of lower weight being used to build up those of higher weight. Checks were obtained by multiplication of factors in different orders.

The writer is indebted to Mr Charles N. Anderson for his valuable aid in the construction of these Tables.

Work on these Tables was supported by Office of Naval Research Contract Nonr 870(00).

REFERENCES

David, F. N. & Kendall, M. G. (1949). *Biometrika*, **36**, 431.
David, F. N. & Kendall, M. G. (1951). *Biometrika*, **38**, 435.
David, F. N. & Kendall, M. G. (1953). *Biometrika*, **40**, 427.
David, F. N. & Kendall, M. G. (1955). *Biometrika*, **42**, 223.

Table 2

$w=2$	(2)	(1^2)
$(1)(1)$	1	2

Table 3

$w=3$	(3)	(21)	(1^3)
$(2)(1)$	1	1	.
$(1^2)(1)$	.	1	3
$(1)^3$	1	3	6

Table 4

$w=4$	(4)	(31)	(2^2)	(21^2)	(1^4)
$(3)(1)$	1	1	.	.	.
$(21)(1)$	.	1	2	2	.
$(2)^2$	1	.	2	.	.
$(2)(1^2)$	.	1	.	1	.
$(1^2)(1)$	.	.	.	1	4
$(1^2)^2$	.	.	1	2	6
$(2)(1)^2$	1	2	2	2	.
$(1^2)(1)^2$	.	1	2	5	12
$(1)^4$	1	4	6	12	24

Table 5

$w=5$	(5)	(41)	(32)	(31^2)	(2^21)	(21^3)	(1^5)
$(4)(1)$	1	1	.	.	.	.	.
$(31)(1)$	.	1	1	2	.	.	.
$(3)(2)$	1	.	1	.	.	.	.
$(3)(1^2)$	.	1	.	1	.	.	.
$(2^2)(1)$	.	.	1	.	1	.	.
$(21^3)(1)$	.	.	.	1	2	3	.
$(21)(2)$	.	1	1	.	2	.	.
$(21)(1^3)$	.	.	1	2	2	3	.
$(2)(1^3)$	.	.	.	1	.	1	.
$(1^5)(1)$	.	.	.	.	.	1	5
$(1^3)(1^2)$	.	.	.	.	1	3	10
$(3)(1)^2$	1	2	1	2	.	.	.
$(21)(1)^3$	.	1	3	4	6	6	.
$(2)^2(1)$	1	1	2	.	2	.	.
$(2)(1^3)(1)$	.	1	1	3	2	3	.
$(1^3)(1)^2$	.	.	.	1	2	7	20
$(1^2)^2(1)$	.	.	1	2	5	12	30
$(2)(1)^3$	1	3	4	6	6	6	.
$(1^2)(1)^3$	.	1	3	7	12	27	60
$(1)^5$	1	5	10	20	30	60	120

Table 6

$w=6$	(6)	(51)	(42)	(41^2)	(3^2)	(321)	(31^3)	(2^3)	(2^21^2)	(21^4)	(1^6)
$(5)(1)$	1	1	.	.	.	.	.	.	.	.	.
$(41)(1)$	.	1	1	2	.	.	.	.	.	.	.
$(4)(2)$	1	.	1	.	.	.	.	.	.	.	.
$(4)(1^2)$	.	1	.	1	.	.	.	.	.	.	.
$(32)(1)$	.	.	1	.	2	1	.	.	.	.	.
$(31^2)(1)$	.	.	.	1	.	1	3	.	.	.	.
$(31)(2)$	.	1	.	.	2	1	.	.	.	.	.
$(31)(1^3)$	.	.	1	2	.	1	3	.	.	.	.
$(3)^2$	1	.	.	.	2	.	.	.	.	.	.
$(3)(21)$	.	1	1	.	.	1	.	.	.	.	.
$(3)(1^3)$	.	.	.	1	.	.	1	.	.	.	.
$(2^21)(1)$	.	.	.	.	.	1	.	3	2	.	.
$(2^3)(2)$	.	.	1	.	.	.	.	3	.	.	.
$(2^3)(1^2)$	.	.	.	.	1	1	.	.	1	.	.
$(21^3)(1)$	.	.	.	.	.	.	1	.	2	4	.
$(21^3)(2)$	.	.	.	1	.	.	.	1	2	.	.
$(21^3)(1^2)$	.	.	.	.	.	1	3	3	4	6	.
$(21)^3$	.	.	1	2	2	2	.	6	4	.	.
$(21)(1^3)$	.	.	.	.	.	1	3	.	2	4	.
$(2)(1^4)$	.	.	.	.	.	.	1	.	.	1	.
$(1^5)(1)$	.	.	.	.	.	.	.	.	.	1	6
$(1^4)(1^2)$	.	.	.	.	.	.	.	.	1	4	15
$(1^3)^2$	.	.	.	.	.	.	.	1	2	6	20
$(4)(1)^2$	1	2	1	2	.	.	.	.	.	.	.
$(31)(1)^3$	.	1	2	4	2	3	6	.	.	.	.
$(3)(2)(1)$	1	1	1	.	2	1	.	.	.	.	.
$(3)(1^3)(1)$	.	1	1	3	.	1	3	.	.	.	.
$(2^3)(1)^2$	.	.	1	.	2	2	.	3	2	.	.
$(21^3)(1)^2$	.	.	.	1	.	3	6	6	10	12	.
$(21)(2)(1)$	.	1	2	2	2	3	.	6	4	.	.
$(21)(1^3)(1)$	.	.	1	2	2	5	9	6	10	12	.
$(2)^3$	1	.	3	.	.	.	.	6	.	.	.
$(2)^3(1^2)$	.	1	.	1	2	2	.	.	2	.	.
$(2)(1^3)(1)$	.	.	1	2	.	2	6	.	4	6	.
$(2)(1^3)^2$	.	.	.	.	.	1	4	3	4	.	.
$(1^4)(1)^2$	.	.	.	.	.	.	1	.	2	9	30
$(1^3)(1^2)(1)$	.	.	.	.	.	1	3	3	8	22	90
$(1^3)^3$	.	.	.	.	1	3	6	6	15	36	90
$(3)(1)^3$	1	3	3	6	2	3	6	.	.	.	.
$(21)(1)^3$	.	1	4	6	6	13	18	18	24	24	.
$(2)^2(1)^2$	1	2	3	2	4	4	.	6	4	.	.
$(2)(1^2)(1)^2$	.	1	2	5	2	6	12	6	10	12	.
$(1^3)(1)^3$	.	.	.	1	.	3	10	6	18	48	120
$(1^2)^2(1)^2$	.	.	1	2	2	8	18	15	34	78	180
$(2)(1)^4$	1	4	7	12	8	16	24	18	24	24	.
$(1^2)(1)^4$	.	1	4	9	6	22	48	36	78	168	360
$(1)^6$	1	6	15	30	20	60	120	90	180	360	720

$w = 7$ (i)	(7)	(61)	(52)	(51^2)	(43)	(421)	(41^3)	(3^21)	(32^2)	(321^3)	(31^4)	(2^31)	(2^21^3)	(21^5)	(1^7)
(6) (1)	1	1	.	.	.	.	.	.	.	.	.	.	.	.	.
(51) (1)	.	1	1	2	.	.	.	.	.	.	.	.	.	.	.
(5) (2)	1	.	1	.	.	.	.	.	.	.	.	.	.	.	.
(5) (1^2)	.	1	.	1	.	.	.	.	.	.	.	.	.	.	.
(42) (1)	.	.	1	.	1	1	.	.	.	.	.	.	.	.	.
(41^2) (1)	.	.	.	1	.	1	3	.	.	.	.	.	.	.	.
(41) (2)	.	1	.	.	1	1	.	.	.	.	.	.	.	.	.
(41) (1^3)	.	.	1	2	.	1	3	.	.	.	.	.	.	.	.
(4) (3)	1	.	.	.	1	.	.	.	.	.	.	.	.	.	.
(4) (21)	.	1	1	.	.	1	.	.	.	.	.	.	.	.	.
(4) (1^3)	.	.	.	1	.	.	1	.	.	.	.	.	.	.	.
(3^2) (1)	.	.	.	.	1	.	.	1	.	.	.	.	.	.	.
(321) (1)	.	.	.	.	.	1	.	2	2	2	.	.	.	.	.
(32) (2)	.	.	1	.	1	.	.	.	2	.	.	.	.	.	.
(32) (1^3)	.	.	.	.	1	1	.	2	.	1	.	.	.	.	.
(31^3) (1)	.	.	.	.	.	.	1	.	.	1	4	.	.	.	.
(31^3) (2)	.	.	.	1	.	.	.	2	.	1	.	.	.	.	.
(31^3) (1^3)	.	.	.	.	.	1	3	.	1	2	6	.	.	.	.
(31) (3)	.	1	.	1	1	.	.	2	.	.	.	.	.	.	.
(31) (21)	.	.	1	2	1	1	.	2	2	2	.	.	.	.	.
(31) (1^3)	.	.	.	.	.	1	3	.	.	1	4	.	.	.	.
(3) (2^2)	.	.	1	.	.	.	.	.	1	.	.	.	.	.	.
(3) (21^3)	.	.	.	1	.	1	.	.	.	1	.	.	.	.	.
(3) (1^4)	.	.	.	.	.	.	1	.	.	.	1	.	.	.	.
(2^31) (1)	.	.	.	.	.	.	.	.	1	.	.	1	.	.	.
(2^21^2) (1)	.	.	.	.	.	1	.	.	1	.	.	3	3	.	.
(2^21) (2)	.	.	.	.	.	1	.	.	1	.	.	3	.	.	.
(2^21) (1^3)	.	.	.	.	.	.	.	1	2	2	.	3	3	.	.
(2^2) (21)	.	.	.	.	1	1	.	.	2	.	.	3	.	.	.
(2^2) (1^3)	.	.	.	.	.	.	.	1	.	1	.	.	1	.	.
(21^4) (1)	.	.	.	.	.	.	.	.	.	.	1	.	2	5	.
(21^4) (2)	.	.	.	.	.	.	1	.	.	1	.	.	2	.	.
(21^4) (1^3)	.	.	.	.	.	.	.	.	.	1	4	3	6	10	.
(21^3) (21)	.	.	.	.	.	1	3	2	2	3	.	6	6	.	.
(21^3) (1^3)	.	.	.	.	.	.	.	.	1	2	6	3	6	10	.
(21) (1^6)	.	.	.	.	.	.	.	.	.	1	4	.	2	5	.
(2) (1^5)	.	.	.	.	.	.	.	.	.	.	1	.	.	1	.
(1^8) (1)	.	.	.	.	.	.	.	.	.	.	.	.	.	1	7
(1^8) (1^5)	.	.	.	.	.	.	.	.	.	.	.	.	1	5	21
(1^8) (1^3)	.	.	.	.	.	.	.	.	.	.	.	1	3	10	35
(5) $(1)^2$	1	2	1	2	.	.	.	.	.	.	.	.	.	.	.
(41) $(1)^2$	.	1	2	4	1	3	6	.	.	.	.	.	.	.	.
(4) (2) (1)	1	1	1	.	1	1	.	.	.	.	.	.	.	.	.
(4) (1^3) (1)	.	1	1	3	.	1	3	.	.	.	.	.	.	.	.
(32) $(1)^2$	.	.	1	.	3	2	.	4	2	2	.	.	.	.	.
(31^3) $(1)^2$	.	.	.	1	.	2	6	2	2	5	12	.	.	.	.
(31) (2) (1)	.	1	1	2	2	1	.	4	2	2	.	.	.	.	.
(31) (1^3) (1)	.	.	1	2	1	4	9	2	2	5	12	.	.	.	.
(3^2) (1)	1	1	.	.	2	.	.	2	.	.	.	.	.	.	.
(3) (21) (1)	.	1	2	2	1	2	.	2	2	2	.	.	.	.	.

$w = 7$ (ii)	(7)	(61)	(52)	(51^2)	(43)	(421)	(41^3)	(3^21)	(32^2)	(321^3)	(31^4)	(2^31)	(2^21^3)	(21^5)	(1^7)
(3) $(2)^2$	1	.	2	.	1	.	.	.	2	.	.	.	.	.	.
(3) (2) (1^3)	.	1	.	1	1	1	.	2	.	1	.	.	.	.	.
(3) (1^3) (1)	.	.	.	1	.	1	4	.	.	1	4	.	.	.	.
$(3)(1^3)^2$	.	.	1	2	.	2	6	.	1	2	6	.	.	.	.
(2^21) $(1)^2$	.	.	.	.	.	1	.	2	5	4	.	9	6	.	.
(2^3) (2) (1)	.	.	1	.	1	1	.	.	3	.	.	3	.	.	.
(2^3) (1^3) (1)	.	.	.	.	1	1	.	3	2	3	.	3	3	.	.
(21^6) $(1)^2$	.	.	.	.	.	.	1	.	.	3	8	6	14	20	.
(21^5) (2) (1)	.	.	.	1	.	2	3	2	2	4	.	6	6	.	.
(21^5) (1^3) (1)	.	.	.	.	.	1	3	2	5	9	18	15	24	30	.
$(21)^3$ (1)	.	.	1	2	3	5	6	6	10	8	.	18	12	.	.
$(21)(2)^2$	.	1	1	.	2	3	.	.	4	.	.	6	.	.	.
(21) (2) (1^3)	.	.	1	2	1	2	3	4	4	5	.	6	6	.	.
(21) (1^5) (1)	.	.	.	.	.	1	3	2	2	7	16	6	14	20	.
(21) $(1^3)^2$	.	.	.	.	1	3	6	4	6	12	24	15	24	30	.
$(2)^2$ (1^3)	.	.	.	1	.	.	1	2	.	2	.	.	2	.	.
(2) (1^5) (1)	.	.	.	.	.	.	1	.	.	1	5	.	2	5	.
(2) (1^8) (1^3)	.	.	.	.	.	1	3	.	1	3	10	3	6	10	.
(1^9) $(1)^2$	.	.	.	.	.	.	.	.	.	.	1	.	2	11	42
(1^6) (1^3) (1)	.	.	.	.	.	.	.	.	.	1	4	3	11	35	105
$(1^3)^2$ (1)	.	.	.	.	.	.	.	.	1	2	6	7	18	50	140
(1^3) $(1^3)^2$	.	.	.	.	.	.	.	1	2	5	12	12	31	80	210
(4) $(1)^3$	1	3	3	6	1	3	6	.	.	.	.	.	.	.	.
(31) $(1)^3$	.	1	3	6	4	9	18	8	6	12	24	.	.	.	.
(3) (2) $(1)^3$	1	2	2	2	3	2	.	4	2	2	.	.	.	.	.
(3) (1^3) $(1)^3$	.	1	2	5	1	5	12	2	2	5	12	.	.	.	.
(2^3) $(1)^3$	.	.	1	.	3	3	.	6	7	6	.	9	6	.	.
(21^3) $(1)^3$	.	.	.	1	.	4	9	6	12	22	36	36	54	60	.
(21) (2) $(1)^3$	.	1	3	4	4	7	6	8	12	10	.	18	12	.	.
(21) (1^3) $(1)^3$	.	.	1	2	3	8	15	12	16	29	48	36	54	60	.
$(2)^4$ (1)	1	1	3	.	3	3	.	.	6	.	.	6	.	.	.
$(2)^3$ (1^3) (1)	.	1	1	3	2	3	3	6	4	6	.	6	6	.	.
(2) (1^3) $(1)^3$	.	.	.	1	.	2	7	2	4	8	20	6	14	20	.
(2) $(1^3)^2$ (1)	.	.	1	2	1	5	12	4	7	14	30	15	24	30	.
(1^9) $(1)^3$	.	.	.	.	.	.	1	.	.	3	13	6	24	75	210
(1^9) (1^3) $(1)^3$	.	.	.	.	.	1	3	2	5	13	34	27	68	170	420
$(1^3)^3$ (1)	.	.	.	.	1	3	6	7	12	27	60	51	117	270	630
$(3)(1)^4$	1	4	6	12	5	12	24	8	6	12	24	.	.	.	.
$(2)^2$ $(1)^3$	1	3	5	6	7	9	6	12	14	12	.	18	12	.	.
(2) $(1)^4$	.	1	5	8	10	23	36	32	44	68	96	90	120	120	.
(2) (1^3) $(1)^3$	.	1	3	7	4	13	27	14	18	34	60	36	54	60	.
(1^5) $(1)^4$	.	.	.	1	.	4	13	6	12	34	88	60	150	360	840
$(1^5)^2$ $(1)^3$	.	.	1	2	3	11	24	18	31	68	150	117	258	570	1260
(2) $(1)^5$	1	5	11	20	15	35	60	40	50	80	120	90	120	120	.
(1^5) $(1)^5$	.	1	5	11	10	35	75	50	80	170	360	270	570	1200	2520
$(1)^7$	1	7	21	42	35	105	210	140	210	420	840	630	1260	2520	5040

Table 8

$w = 8$ (i)	(8)	(71)	(62)	(61^2)	(53)	(521)	(51^3)	(4^2)	(431)	(42^2)	(421^2)
(7) (1)	1	1	.	.	.	.	.	.	.	.	.
(61) (1)	.	1	1	2	.	.	.	.	.	.	.
(6) (2)	1	.	1	.	.	.	.	.	.	.	.
$(6)(1^2)$	.	1	.	1	.	.	.	.	.	.	.
(52) (1)	.	.	1	.	1	.	.	.	.	.	.
(51^3) (1)	.	.	.	1	.	1	.	.	.	.	.
(51) (2)	.	1	.	.	1	1	3	.	.	.	.
$(51)(1^2)$	.	.	1	2	.	1	3	.	.	.	.
(5) (3)	1	.	.	.	1	.	.	.	.	.	.
(5) (21)	.	1	1	.	.	1	.	.	.	.	.
$(5)(1^3)$	.	.	.	1	.	.	1	.	.	.	.
(43) (1)	.	.	.	.	1	.	.	2	1	.	.
(421) (1)	.	.	.	.	.	1	.	.	1	2	2
(42) (2)	.	.	1	.	.	.	.	2	.	2	.
$(42)(1^2)$	.	.	.	.	1	1	.	.	1	.	1
(41^3) (1)	.	.	.	1	.	.	1	.	.	.	1
(41^3) (2)	.	.	.	.	.	1	.	.	1	.	1
$(41^2)(1^2)$	.	.	.	.	1	.	3	.	.	1	2
(41) (3)	.	1	.	.	.	.	.	2	1	2	.
(41) (21)	.	.	1	2	1	1	.	.	1	2	2
$(41)(1^3)$	.	.	.	.	.	1	3	.	.	.	1
$(4)^2$	1	.	.	.	.	.	.	2	.	.	.
(4) (31)	.	1	.	.	1	.	.	.	1	.	.
$(4)(2^2)$	.	.	1	.	.	.	.	.	.	1	.
$(4)(21^2)$	.	.	.	1	.	1	.	.	.	.	1
$(4)(1^4)$	.	.	.	.	.	.	1	.	.	.	.
(3^21) (1)	.	.	.	.	1	.	.	.	1	.	.
(3^2) (2)	.	.	.	.	.	.	.	.	.	1	.
$(3^2)(1^2)$	.	.	.	.	.	.	.	1	1	.	.
(32^2) (1)	.	.	.	.	.	.	.	.	.	1	.
(321^3) (1)	.	.	.	.	.	.	.	.	.	.	1
(321) (2)	.	.	.	.	.	1	.	.	1	.	.
$(321)(1^3)$	.	.	.	.	.	.	.	.	1	2	2
(32) (3)	.	.	1	.	1	.	.	2	1	2	.
(32) (21)	.	.	.	.	1	1	.	.	1	.	1
$(32)(1^3)$	.	.	.	.	.	.	.	1	.	.	.
(31^5) (1)	.	.	.	.	.	.	1	.	.	.	.
(31^5) (2)	.	.	.	.	.	.	.	.	1	.	1
(31^5) (1^5)	.	.	.	1	.	.	.	.	.	.	.
(31^5) (3)	.	.	.	.	.	.	.	.	1	.	1
(31^3) (21)	.	.	.	.	.	1	3	.	1	.	1
(31^3) (1^5)	.	.	.	.	.	.	.	.	.	1	2
$(31)^2$	.	.	1	2	.	.	.	2	2	.	.
(31) (2^2)	.	.	.	.	1	1	.	.	.	.	.
(31) (21^3)	.	.	.	.	.	1	3	.	1	2	2
(31) (1^4)	.	.	.	.	.	.	.	.	.	.	1
$(3)(2^21)$	.	.	.	.	.	1	.	.	.	1	.
$(3)(21^3)$	.	.	.	.	.	.	1	.	.	.	1
$(3)(1^5)$	.	.	.	.	.	.	.	.	.	.	.
(2^21) (1)	.	.	.	.	.	.	.	.	.	1	.
(2^2) (2)	.	.	.	.	.	.	.	.	.	.	.
$(2^2)(1^2)$	.	.	.	.	.	.	.	.	.	.	.
(2^21^3) (1)	.	.	.	.	.	.	.	.	.	.	1
(2^21^3) (2)	.	.	.	.	.	.	.	.	.	.	.
(2^21^3) (1^2)	.	.	.	.	.	.	.	.	.	.	.
(2^21) (21)	.	.	.	.	.	.	.	.	1	2	2
(2^21) (1^3)	.	.	.	.	.	.	.	.	.	.	.
$(2^2)^2$	.	.	.	.	.	.	.	1	.	2	.
(2^2) (21^2)	.	.	.	.	.	.	.	.	1	.	1
(2^2) (1^4)	.	.	.	.	.	.	.	.	.	.	.
(21^5) (1)	.	.	.	.	.	.	.	.	.	.	.
(21^4) (2)	.	.	.	.	.	.	.	.	.	.	.
(21^4) (1^2)	.	.	.	.	.	.	.	.	.	.	.
(21^3) (21)	.	.	.	.	.	.	.	.	.	.	1
(21^3) (1^3)	.	.	.	.	.	.	.	.	.	.	.
$(21^3)^2$	.	.	.	.	.	.	.	.	.	1	2
(21^3) (1^4)	.	.	.	.	.	.	.	.	.	.	.
(21) (1^5)	.	.	.	.	.	.	.	.	.	.	.
(2) (1^6)	.	.	.	.	.	.	.	.	.	.	.
(1^7) (1)	.	.	.	.	.	.	.	.	.	.	.
(1^6) (1^2)	.	.	.	.	.	.	.	.	.	.	.
(1^5) (1^3)	.	.	.	.	.	.	.	.	.	.	.
$(1^4)^2$	.	.	.	.	.	.	.	.	.	.	.

Table 8 (*cont.*)

$w = 8$ (ii)	$(4^2 1^4)$	$(3^2 2)$	$(3^2 1^2)$	$(32^2 1)$	(321^3)	(31^5)	(2^4)	$(2^3 1^2)$	$(2^2 1^4)$	(21^6)	(1^8)
(7) (1)	.	.	.	.	.	.	.	.	.	.	.
(61) (1)	.	.	.	.	.	.	.	.	.	.	.
(6) (2)	.	.	.	.	.	.	.	.	.	.	.
(6) (1^2)	.	.	.	.	.	.	.	.	.	.	.
(52) (1)	.	.	.	.	.	.	.	.	.	.	.
(51^3) (1)	.	.	.	.	.	.	.	.	.	.	.
(51) (2)	.	.	.	.	.	.	.	.	.	.	.
(51) (1^2)	.	.	.	.	.	.	.	.	.	.	.
(5) (3)	.	.	.	.	.	.	.	.	.	.	.
(5) (21)	.	.	.	.	.	.	.	.	.	.	.
(5) (1^3)	.	.	.	.	.	.	.	.	.	.	.
(43) (1)	.	.	.	.	.	.	.	.	.	.	.
(421) (1)	.	.	.	.	.	.	.	.	.	.	.
(42) (2)	.	.	.	.	.	.	.	.	.	.	.
(42) (1^2)	.	.	.	.	.	.	.	.	.	.	.
(41^3) (1)	4	.	.	.	.	.	.	.	.	.	.
(41^3) (2)	.	.	.	.	.	.	.	.	.	.	.
(41^3) (1^2)	6	.	.	.	.	.	.	.	.	.	.
(41) (3)	.	.	.	.	.	.	.	.	.	.	.
(41) (21)	.	.	.	.	.	.	.	.	.	.	.
(41) (1^3)	4	.	.	.	.	.	.	.	.	.	.
$(4)^2$	.	.	.	.	.	.	.	.	.	.	.
(4) (31)	.	.	.	.	.	.	.	.	.	.	.
(4) (2^2)	.	.	.	.	.	.	.	.	.	.	.
(4) (21^2)	.	.	.	.	.	.	.	.	.	.	.
(4) (1^4)	1	.	.	.	.	.	.	.	.	.	.
$(3^2 1)$ (1)	.	1	2	.	.	.	.	.	.	.	.
(3^2) (2)	.	1	.	.	.	.	.	.	.	.	.
(3^2) (1^2)	.	.	1	.	.	.	.	.	.	.	.
(32^2) (1)	.	2	.	1	.	.	.	.	.	.	.
(321^3) (1)	.	.	2	2	3	.	.	.	.	.	.
(321) (2)	.	2	.	2	.	.	.	.	.	.	.
(321) (1^3)	.	2	4	2	3	.	.	.	.	.	.
(32) (3)	.	2	.	.	.	.	.	.	.	.	.
(32) (21)	.	2	.	2	.	.	.	.	.	.	.
(32) (1^3)	.	.	2	.	1	.	.	.	.	.	.
(31^5) (1)	1	.	.	.	1	5	.	.	.	.	.
(31^5) (2)	.	.	2	.	1	.	.	.	.	.	.
(31^5) (1^2)	4	.	.	1	3	10	.	.	.	.	.
(31^5) (3)	.	.	2	.	.	.	.	.	.	.	.
(31^5) (21)	.	2	4	2	3	.	.	.	.	.	.
(31^5) (1^3)	6	.	.	1	3	10	.	.	.	.	.
$(31)^2$	.	2	4	.	.	.	.	.	.	.	.
(31) (2^2)	.	2	.	1	.	.	.	.	.	.	.
(31) (21^4)	.	.	2	2	3	.	.	.	.	.	.
(31) (1^4)	4	.	.	.	1	5	.	.	.	.	.
(3) $(2^2 1)$	.	.	.	1	.	.	.	.	.	.	.
(3) (21^5)	.	.	.	.	1	.	.	.	.	.	.
(3) (1^6)	1	.	.	.	.	1	.	.	.	.	.
$(2^3 1)$ (1)	.	.	.	1	.	.	4	2	.	.	.
(2^3) (2)	.	.	.	.	.	.	4	.	.	.	.
(2^3) (1^2)	.	1	.	1	.	.	.	1	.	.	.
$(2^3 1^2)$ (1)	.	.	.	.	1	.	.	3	4	.	.
$(2^3 1^2)$ (2)	.	.	.	1	.	.	.	3	.	.	.
$(2^3 1^2)$ (1^3)	.	.	1	2	3	.	6	6	6	.	.
$(2^3 1)$ (21)	.	2	.	3	.	.	12	6	.	.	.
$(2^3 1)$ (1^3)	.	1	2	2	3	.	.	3	4	.	.
$(2^3)^2$	.	.	.	.	.	.	6	.	.	.	.
(2^5) (21^3)	.	1	.	2	.	.	.	3	.	.	.
(2^5) (1^4)	.	.	1	.	1	.	.	.	1	.	.
(21^6) (1)	.	.	.	.	.	1	.	.	2	6	.
(21^6) (2)	1	.	.	.	1	.	.	.	2	.	.
(21^6) (1^2)	.	.	.	.	1	5	.	3	8	15	.
(21^6) (21)	4	.	2	2	4	.	.	6	8	.	.
(21^6) (1^3)	.	.	.	1	3	10	4	6	12	20	.
$(21^3)^2$	6	2	4	4	6	.	12	12	12	.	.
(21^3) (1^4)	.	.	.	1	3	10	.	3	8	15	.
(21) (1^5)	.	.	.	.	1	5	.	.	2	6	.
(2) (1^6)	.	.	.	.	.	1	.	.	.	1	.
(1^8) (1)	.	.	.	.	.	.	.	.	.	1	8
(1^8) (1^2)	.	.	.	.	.	.	.	.	1	6	28
(1^6) (1^2)	.	.	.	.	.	.	.	1	4	15	56
$(1^4)^2$	.	.	.	.	.	.	1	2	6	20	70

Table 8 (*cont.*)

$w = 8$ (iii)	(8)	(71)	(62)	(61^2)	(53)	(521)	(51^3)	(4^2)	(431)	(42^2)	(421^2)
$(6)(1)^2$	1	2	1	2	.	.	.	.	.	.	.
$(51)(1)^3$	.	1	2	4	1	3	6	.	.	.	.
$(5)(2)(1)$	1	1	1	.	1	1	.	.	.	.	.
$(5)(1^3)(1)$	.	1	1	3	.	1	3	.	.	.	.
$(42)(1)^2$	.	.	1	.	2	2	.	2	2	2	2
$(41^2)(1)^3$	.	.	.	1	.	2	6	.	1	2	5
$(41)(2)(1)$	.	1	1	2	1	1	.	2	2	2	2
$(41)(1^3)(1)$	.	.	1	2	1	4	9	.	1	2	5
$(4)(3)(1)$	1	1	.	.	1	.	.	2	1	.	.
$(4)(21)(1)$	.	1	2	2	1	2	.	.	1	2	2
$(4)(2)^2$	1	.	2	.	.	.	.	2	.	2	.
$(4)(2)(1^2)$	.	1	.	1	1	1	.	.	1	.	1
$(4)(1^4)(1)$	.	.	.	1	.	1	4	.	.	.	1
$(4)(1^2)^2$	.	.	1	2	.	2	6	.	.	1	2
$(3^2)(1)^2$	.	.	.	.	1	.	.	2	2	.	.
$(321)(1)^3$	.	.	.	.	.	1	.	.	3	4	4
$(32)(2)(1)$	.	.	1	.	2	1	.	2	1	2	.
$(32)(1^3)(1)$	.	.	.	.	1	1	.	2	4	2	3
$(31^2)(1)^3$	.	.	.	.	.	.	1	.	.	.	2
$(31^2)(2)(1)$	.	.	.	1	.	1	3	.	2	.	1
$(31^2)(1^2)(1)$	.	.	.	.	.	1	3	.	1	3	7
$(31)(3)(1)$	.	1	1	2	1	.	.	2	3	.	.
$(31)(21)(1)$	.	.	1	2	2	4	6	2	4	4	4
$(31)(2)^2$	.	1	.	.	3	2	.	.	1	.	.
$(31)(2)(1^2)$	.	.	1	2	.	1	3	2	3	2	2
$(31)(1^4)(1)$	.	.	.	.	.	1	3	.	1	2	6
$(31)(1^2)^2$	.	.	.	.	1	3	6	.	2	4	10
$(3)^2(2)$	1	.	1	.	2	.	.	.	.	.	.
$(3)^2(1^2)$	.	1	.	1	.	.	.	2	2	.	.
$(3)(2^2)(1)$	.	.	1	.	1	1	.	.	.	1	.
$(3)(21^3)(1)$	.	.	.	1	.	2	3	.	1	2	3
$(3)(21)(2)$	.	1	1	.	1	2	.	2	1	2	.
$(3)(21)(1^3)$	.	.	1	2	1	2	3	.	2	2	3
$(3)(2)(1^3)$	.	.	.	1	.	.	1	.	1	.	1
$(3)(1^5)(1)$	.	.	.	.	.	.	1	.	.	.	1
$(3)(1^3)(1^2)$	.	.	.	.	.	1	3	.	.	1	3
$(2^3)(1)^2$	.	.	.	.	.	.	.	.	.	1	.
$(2^2)(1^4)(1)^2$	.	.	.	.	.	.	.	.	.	.	1
$(2^2)(2)(1)$	.	.	.	.	.	1	.	.	1	3	2
$(2^2)(1^3)(1)$	.	.	.	.	.	.	.	.	1	2	2
$(2^2)(21)(1)$	.	.	.	.	1	1	.	2	2	4	2
$(2^2)(2)^2$	.	.	1	.	.	.	.	2	.	5	.
$(2^2)(2)(1^2)$	.	.	.	.	1	1	.	.	1	.	1
$(2^2)(1^4)(1)$	.	.	.	.	.	.	.	.	1	.	1
$(2^2)(1^2)^2$	.	.	.	.	.	.	.	1	2	2	2
$(21^6)(1)^2$	.	.	.	.	.	.	.	.	.	.	.
$(21^5)(2)(1)$	.	.	.	.	.	.	1	.	.	.	2
$(21^4)(1^3)(1)$	.	.	.	.	.	.	.	.	.	.	1
$(21^3)(21)(1)$	.	.	.	.	.	1	3	.	3	4	8
$(21^4)(2)^2$	.	.	.	1	.	1	.	.	2	.	3
$(21^3)(2)(1^2)$	.	.	.	.	.	1	3	.	1	3	4
$(21^3)(1^3)(1)$	.	.	.	.	.	.	.	.	.	1	2
$(21^2)(1^3)^2$	.	.	.	.	.	.	.	.	1	2	5
$(21)^4(2)$	.	.	1	2	2	2	.	2	4	8	6
$(21)^4(1^2)$	.	.	.	.	1	3	6	2	5	6	9
$(21)(2)(1^3)$	.	.	.	.	.	1	3	.	1	.	2
$(21)(1^5)(1)$	.	.	.	.	.	.	.	.	.	.	1
$(21)(1^3)(1^2)$	.	.	.	.	.	.	.	.	1	2	5
$(2)^3(1^4)$	.	.	.	.	.	.	1	.	.	.	.
$(2)(1^3)(1)$	.	.	.	.	.	.	.	.	.	.	.
$(2)(1^4)(1^2)$	.	.	.	.	.	.	.	.	.	.	1
$(2)(1^2)^3$	.	.	.	.	.	.	.	.	.	1	2
$(1^8)(1)^2$	.	.	.	.	.	.	.	.	.	.	.
$(1^6)(1^2)(1)$	.	.	.	.	.	.	.	.	.	.	.
$(1^5)(1^3)(1)$	.	.	.	.	.	.	.	.	.	.	.
$(1^4)(1^2)^2$	.	.	.	.	.	.	.	.	.	.	.
$(1^2)^3(1^2)$	.	.	.	.	.	.	.	.	.	.	.

Table 8 (*cont.*)

$w = 8$ (iv)	(41^4)	(3^22)	(3^21^2)	(32^21)	(321^3)	(31^5)	(2^4)	(2^31^2)	(2^21^4)	(21^6)	(1^8)
$(6)(1)^2$	.	.	.	.	.	.	.	.	.	.	.
$(51)(1)^3$	.	.	.	.	.	.	.	.	.	.	.
$(5)(2)(1)$	.	.	.	.	.	.	.	.	.	.	.
$(5)(1^3)(1)$	.	.	.	.	.	.	.	.	.	.	.
$(42)(1)^4$	.	.	.	.	.	.	.	.	.	.	.
$(41^3)(1)^2$	12	.	.	.	.	.	.	.	.	.	.
$(41)(2)(1)$	.	.	.	.	.	.	.	.	.	.	.
$(41)(1^3)(1)$	12	.	.	.	.	.	.	.	.	.	.
$(4)(3)(1)$	.	.	.	.	.	.	.	.	.	.	.
$(4)(21)(1)$	.	.	.	.	.	.	.	.	.	.	.
$(4)(2)^2$	.	.	.	.	.	.	.	.	.	.	.
$(4)(2)(1^4)$	.	.	.	.	.	.	.	.	.	.	.
$(4)(1^5)(1)$	4	.	.	.	.	.	.	.	.	.	.
$(4)(1^4)^2$	6	.	.	.	.	.	.	.	.	.	.
$(3^3)(1)^2$	.	1	2	.	.	.	.	.	.	.	.
$(321)(1)^3$	.	6	8	6	6	.	.	.	.	.	.
$(32)(2)(1)$	.	4	.	2	.	.	.	.	.	.	.
$(32)(1^4)(1)$	.	2	6	2	3	.	.	.	.	.	.
$(31^4)(1)^3$	8	.	2	2	7	20	.	.	.	.	.
$(31^4)(2)(1)$	.	2	6	2	3	.	.	.	.	.	.
$(31^3)(1^3)(1)$	18	2	4	5	12	30	.	.	.	.	.
$(31)(3)(1)$	.	2	4	.	.	.	.	.	.	.	.
$(31)(21)^2(1)$	.	6	8	6	6	.	.	.	.	.	.
$(31)(2)^2$	.	4	.	2	.	.	.	.	.	.	.
$(31)(2)(1^4)$	.	2	6	2	3	.	.	.	.	.	.
$(31)(1^5)(1)$	16	.	2	2	7	20	.	.	.	.	.
$(31)(1^4)^2$	24	2	4	5	12	30	.	.	.	.	.
$(3)^2(2)$	.	2	.	.	.	.	.	.	.	.	.
$(3)^2(1^2)$	.	.	2	.	.	.	.	.	.	.	.
$(3)(2^3)(1)$	.	2	.	1	.	.	.	.	.	.	.
$(3)(21^3)(1)$	.	.	2	2	3	.	.	.	.	.	.
$(3)(21)(2)$	.	2	.	2	.	.	.	.	.	.	.
$(3)(21)(1^4)$	.	2	4	2	3	.	.	.	.	.	.
$(3)(2)(1^3)$	.	.	2	.	1	.	.	.	.	.	.
$(3)(1^6)(1)$	5	.	.	.	1	5	.	.	.	.	.
$(3)(1^5)(1^3)$	10	.	.	1	3	10	.	.	.	.	.
$(2^5)(1)^3$	.	2	.	2	.	.	4	2	.	.	.
$(2^31^2)(1)^2$	.	.	2	5	6	.	12	15	12	.	.
$(2^31)(2)(1)$	.	2	.	4	.	.	12	6	.	.	.
$(2^31)(1^4)(1)$	.	5	6	9	9	.	12	15	12	.	.
$(2^3)(21)(1)$	.	4	.	5	.	.	12	6	.	.	.
$(2^3)(2)^2$	.	.	.	.	.	.	12	.	.	.	.
$(2^5)(2)(1^2)$	.	3	.	3	.	.	.	3	.	.	.
$(2^5)(1^3)(1)$	.	1	4	2	4	.	.	3	4	.	.
$(2^4)(1^2)^2$	.	2	6	4	6	.	6	6	6	.	.
$(21^6)(1)^2$	1	.	.	.	3	10	.	6	18	30	.
$(21^5)(2)(1)$	4	.	2	2	5	.	.	6	8	.	.
$(21^5)(1^3)(1)$	4	.	2	5	13	30	12	24	44	60	.
$(21^5)(21)(1)$	12	6	10	14	15	.	24	30	24	.	.
$(21^5)(2)^2$	.	2	.	4	.	.	.	6	.	.	.
$(21^4)(2)(1^3)$	6	2	6	6	9	.	12	12	12	.	.
$(21^4)(1^5)(1)$	6	2	4	8	18	40	12	24	44	60	.
$(21^4)(1^3)^2$	12	5	8	16	30	60	24	45	72	90	.
$(21)^3(2)$	.	6	.	8	.	.	24	12	.	.	.
$(21)^3(1^2)$	12	10	14	18	18	.	24	30	24	.	.
$(21)(2)(1^5)$	4	2	6	4	7	.	.	6	8	.	.
$(21)(1^6)(1)$	4	.	2	2	9	25	.	6	18	30	.
$(21)(1^5)(1^3)$	12	2	6	9	22	50	12	24	44	60	.
$(2)^3(1^4)$	1	.	2	.	2	.	.	.	2	.	.
$(2)(1^5)(1)$	1	.	.	.	1	6	.	.	2	6	.
$(2)(1^6)(1^2)$	4	.	.	1	4	15	.	3	8	15	.
$(2)(1^5)^2$	6	.	.	2	6	20	4	6	12	20	.
$(1^8)(1)^2$	.	.	.	.	.	1	.	.	2	13	56
$(1^8)(1^5)(1)$	.	.	.	.	1	5	.	3	14	51	168
$(1^6)(1^5)(1)$	.	.	.	1	3	10	4	11	32	95	280
$(1^6)(1^3)^2$	.	.	1	2	7	20	6	18	53	150	420
$(1^5)^2(1^2)$	.	1	2	5	12	30	12	31	80	210	560

Table 8 (*cont.*)

$w = 8\ (v)$	(8)	(71)	(62)	(51^2)	(53)	(521)	(51^3)	(4^2)	(431)	(42^2)	(421^2)
$(5)\,(1)^3$	1	3	3	6	1	3	6	.	.	.	.
$(41)\,(1)^3$	.	1	3	6	3	9	18	2	4	6	12
$(4)\,(2)\,(1)^2$	1	2	2	2	2	2	.	2	2	2	2
$(4)\,(1^3)\,(1)^2$	.	1	2	5	1	5	12	.	1	2	5
$(32)\,(1)^2$	.	.	1	.	4	3	.	6	9	6	6
$(31^3)\,(1)^2$	.	.	.	1	.	3	9	.	4	6	15
$(31)\,(2)\,(1)^2$	.	1	2	4	3	4	6	4	7	4	4
$(31)\,(1^3)\,(1)^2$	.	.	1	2	2	7	15	2	7	10	22
$(3)^2\,(1)^2$	1	2	1	2	2	.	.	4	4	.	.
$(3)\,(21)\,(1)^2$	.	1	3	4	3	6	6	2	5	6	6
$(3)\,(2)^2\,(1)$	1	1	2	.	3	2	.	2	1	2	.
$(3)\,(2)\,(1^4)\,(1)$	.	1	1	3	1	2	3	2	4	2	3
$(3)\,(1^5)\,(1)^2$	.	.	.	1	.	2	7	.	1	2	7
$(3)\,(1^3)^2\,(1)$	.	.	1	2	1	5	12	.	2	5	12
$(2^31)\,(1)^2$	.	.	.	.	.	1	.	.	3	7	6
$(2^3)\,(2)\,(1)^2$	.	.	1	.	2	2	.	2	2	5	2
$(2^3)\,(1^2)\,(1)^2$	.	.	.	.	1	1	.	2	5	4	5
$(21^5)\,(1)^2$	.	.	.	.	.	.	1	.	.	.	4
$(21^4)\,(2)\,(1)^2$	.	.	.	1	.	3	6	.	4	6	11
$(21^3)\,(1^3)\,(1)^2$	.	.	.	.	.	1	3	.	3	7	14
$(21)^2\,(1)^2$	.	.	1	2	4	8	12	6	14	20	24
$(21)\,(2)^2\,(1)$	.	1	2	2	3	4	.	4	5	10	6
$(21)(2)(1^4)(1)$	.	.	1	2	2	5	9	2	7	8	12
$(21)\,(1^5)\,(1)^2$	.	.	.	.	.	1	3	.	3	4	12
$(21)\,(1^3)^2\,(1)$	.	.	.	.	1	3	6	2	8	12	24
$(2)^4$	1	.	4	.	.	.	.	6	.	12	.
$(2)^3\,(1^2)$	.	1	.	1	3	3	.	.	3	.	3
$(2)^2\,(1^4)\,(1)$	.	.	.	1	.	1	4	.	2	.	3
$(2)^2\,(1^3)^2$	.	.	1	2	.	2	6	2	4	5	6
$(2)\,(1^6)\,(1)^2$	.	.	.	.	.	.	1	.	.	.	2
$(2)(1^3)(1^3)(1)$	.	.	.	.	.	1	3	.	1	3	8
$(2)\,(1^4)^2$	.	.	.	.	1	3	6	.	3	6	15
$(1^8)\,(1)^2$	.	.	.	.	.	.	.	.	.	.	.
$(1^6)\,(1^2)\,(1)^2$	.	.	.	.	.	.	.	.	.	.	1
$(1^5)^2\,(1)^2$	.	.	.	.	.	.	.	.	.	1	2
$(1^5)\,(1^3)^2\,(1)$	.	.	.	.	.	.	.	.	1	2	5
$(1^4)^2$	.	.	.	.	.	.	.	1	4	6	12
$(4)\,(1)^4$	1	4	6	12	4	12	24	2	4	6	12
$(31)\,(1)^4$	.	1	4	8	7	18	36	8	21	24	48
$(3)\,(2)\,(1)^3$	1	3	4	6	5	6	6	6	9	6	6
$(3)\,(1^5)\,(1)^3$		1	3	7	3	12	27	2	8	12	27
$(2^31)\,(1)^4$		.	1	.	4	4	.	6	12	13	12
$(21^5)\,(1)^4$	.	.	.	1	.	5	12	.	10	20	39
$(21)\,(2)\,(1)^3$	.	1	4	6	7	14	18	8	19	26	30
$(21)\,(1^3)\,(1)^3$	.	.	1	2	4	11	21	6	23	32	60
$(2)^4\,(1)^2$	1	2	4	2	6	6	.	6	6	12	6
$(2)^3\,(1^2)\,(1)^2$	.	1	2	5	3	7	12	4	11	10	15
$(2)\,(1^7)\,(1)^2$	.	.	.	1	.	3	10	.	4	6	19
$(2)\,(1^5)^2\,(1)^2$	.	.	1	2	2	8	18	2	10	17	36
$(1^8)\,(1)^2$	.	.	.	.	.	.	1	.	.	.	4
$(1^7)\,(1^2)\,(1)^2$	*	.	.	.	.	1	3	.	3	7	18
$(1^6)^2\,(1)^2$	.	.	.	.	1	3	6	2	11	18	30
$(3)\,(1)^5$	1	5	10	20	11	30	60	10	25	30	60
$(21)\,(1)^5$	.	1	6	10	15	36	60	20	65	90	150
$(2)^4\,(1)^4$	1	4	8	12	12	20	24	14	28	32	36
$(2)\,(1^3)\,(1)^4$	.	1	4	9	7	23	48	8	31	44	87
$(1^8)\,(1)^3$	.	.	.	1	.	5	16	.	10	20	55
$(1^6)^2\,(1)^4$	.	.	1	2	4	14	30	6	32	53	114
$(2)\,(1)^6$	1	6	16	30	26	66	120	30	90	120	210
$(1^8)\,(1)^6$	.	1	6	13	15	51	108	20	95	150	315
$(1)^8$	1	8	28	56	56	168	336	70	280	420	840

Table 8 (*cont.*)

$w = 8$ (vi)	(41^4)	(3^22)	(3^212^2)	(32^21)	(321^3)	(31^5)	(2^4)	(2^212^2)	(2^21^4)	(21^6)	(1^8)
$(5)(1)^3$	.	.	.	.	.	.	.	.	.	.	.
$(41)(1)^3$	24	.	.	.	.	.	.	.	.	.	.
$(4)(2)(1)^2$	.	.	.	.	.	.	.	.	.	.	.
$(4)(1^2)(1)^2$	12	.	.	.	.	.	.	.	.	.	.
$(32)(1)^2$	.	8	12	6	6	.	.	.	.	.	.
$(31^2)(1)^2$	36	6	14	12	27	60	.	.	.	.	.
$(31)(2)(1)^2$	.	8	12	6	6	.	.	.	.	.	.
$(31)(1^2)(1)^2$	48	6	14	12	27	60	.	.	.	.	.
$(3)^2(1)^2$	.	2	4	.	.	.	.	.	.	.	.
$(3)(21)(1)^2$	.	6	8	6	6	.	.	.	.	.	.
$(3)(2)^2(1)$	.	4	.	2	.	.	.	.	.	.	.
$(3)(2)(1^2)(1)$	.	2	6	2	3	.	.	.	.	.	.
$(3)(1^3)(1)^2$	20	.	2	2	7	20	.	.	.	.	.
$(3)(1^2)^2(1)$	30	2	4	5	12	30	.	.	.	.	.
$(2^21)(1)^2$	.	12	12	22	18	.	36	36	24	.	.
$(2^2)(2)(1)^2$	.	6	.	6	.	.	12	6	.	.	.
$(2^2)(1^2)(1)^2$	.	7	12	11	12	.	12	15	12	.	.
$(21^3)(1)^2$	12	.	6	12	31	60	24	54	96	120	.
$(21^2)(2)(1)^2$	12	6	12	16	18	.	24	30	24	.	.
$(21^2)(1^2)(1)^2$	30	12	22	38	69	120	60	102	156	180	.
$(21)^2(1)^2$	24	26	28	44	36	.	72	72	48	.	.
$(21)(2)^2(1)$	.	8	.	10	.	.	24	12	.	.	.
$(21)(2)(1^3)(1)$	12	12	18	20	21	.	24	30	24	.	.
$(21)(1^4)(1)^2$	28	6	18	22	51	100	24	54	96	120	.
$(21)(1^2)^2(1)$	48	16	32	45	84	150	60	102	156	180	.
$(2)^4$	.	.	.	.	.	.	24	.	.	.	.
$(2)^3(1^2)$	.	6	.	6	.	.	.	6	.	.	.
$(2)^2(1^2)(1)$	4	2	8	4	8	.	.	6	8	.	.
$(2)^2(1^4)^2$	6	4	12	8	12	.	12	12	12	.	.
$(2)(1^4)(1)^2$	9	.	2	2	10	30	.	6	18	30	.
$(2)(1^3)(1^2)(1)$	22	2	6	10	25	60	12	24	44	60	.
$(2)(1^3)^2$	36	7	12	21	42	90	24	45	72	90	.
$(1^8)(1)^2$	1	.	.	.	3	16	.	6	30	108	336
$(1^6)(1^2)(1)^2$	4	.	2	5	18	55	12	39	114	315	840
$(1^5)^2(1)^2$	6	2	4	12	30	80	28	68	172	440	1120
$(1^6)(1^2)^2(1)$	12	5	12	24	58	140	48	117	284	690	1680
$(1^2)^4$	24	12	28	48	108	240	90	204	468	1080	2520
$(4)(1)^4$	24	.	.	.	.	.	.	.	.	.	.
$(31)(1)^4$	96	20	40	30	60	120	.	.	.	.	.
$(3)(2)(1)^3$	.	8	12	6	6	.	.	.	.	.	.
$(3)(1^2)(1)^3$	60	6	14	12	27	60	.	.	.	.	.
$(2^2)(1)^4$	.	20	24	28	24	.	36	36	24	.	.
$(21^2)(1)^4$	72	30	56	92	156	240	144	234	336	360	.
$(21)(2)(1)^3$	24	32	36	50	42	.	72	72	48	.	.
$(21)(1^3)(1)^3$	108	44	82	110	189	300	144	234	336	360	.
$(2)^3(1)^2$	.	12	.	12	.	.	24	12	.	.	.
$(2)^2(1^2)(1)^2$	12	14	24	22	24	.	24	30	24	.	.
$(2)(1^3)(1)^3$	48	6	20	24	58	120	24	54	96	120	.
$(2)(1^2)^2(1)^2$	78	18	36	50	96	180	60	102	156	180	.
$(1^6)(1)^4$	17	.	6	12	46	140	24	84	246	660	1680
$(1^3)(1^2)(1)^3$	46	12	30	58	141	340	108	258	612	1440	3360
$(1^2)^2(1)^2$	84	31	68	117	258	570	204	453	1008	2250	5040
$(3)(1)^5$	120	20	40	30	60	120	.	.	.	.	.
$(21)(1)^5$	240	120	200	270	420	600	360	540	720	720	.
$(2)^2(1)^4$	24	40	48	56	48	.	72	72	48	.	.
$(2)(1^2)(1)^4$	168	50	96	122	216	360	144	234	336	360	.
$(1^2)(1)^6$	140	30	80	140	340	800	240	570	1320	3000	6720
$(1^2)^2(1)^4$	246	80	172	284	612	1320	468	1008	2172	4680	10080
$(2)(1)^6$	360	140	240	300	480	720	360	540	720	720	.
$(1^2)(1)^6$	660	210	440	690	1440	3000	1080	2250	4680	9720	20160
$(1)^8$	1680	560	1120	1680	3360	6720	2520	5040	10080	20160	40320

MISCELLANEA

Approximate linearization of the incomplete β-function

BY A. W. KIMBALL AND E. LEACH
Oak Ridge National Laboratory, Tennessee†

1. INTRODUCTION

The incomplete β-function has been a favourite subject with statisticians for many years because it arises frequently in our work and because it presents problems in evaluation. Several tables of this function and related functions have been constructed and many approximations have been derived. Wise (1950) gives an excellent account of various approximations in terms of percentage points of the χ^2-distribution and provides methods for computing percentage errors. From his paper, and other literature which is too voluminous to review here, one can ascertain that most efforts have been directed toward the numerical evaluation of either the function itself or of one of its arguments, usually in connexion with the determination of Type I or Type II errors for test statistics, the construction of confidence intervals, or the determination of sample sizes. When one attempts to use existing formulae and tables, one is immediately impressed by the fact that no single table or formula is sufficient for every purpose. Tables may be unsatisfactory either because the range of the arguments is too narrow or because interpolation is required. The approximation formulae may be unsatisfactory because they are derived for significance tests or confidence intervals and hence are restricted to percentage points usually near the tails of the distribution.

The approximation reported in this paper resulted from the need for a single representation of the incomplete β-function that would be valid over wide ranges of the function and the three arguments. In radiation mortality studies, an approach known popularly as 'target theory' predicts that, for some organisms, the probability (M) of death from a dose (D) of radiation is given by

$$M = \sum_{i=p}^{p+q-1} C_i^{p+q-1} x^i (1-x)^{p+q-1-i} = I_x(p, q),$$

where p and q are unknown parameters, x is a function of D, and $I_x(p, q)$ is Pearson's (1934) notation for the incomplete beta function, i.e.

$$I_x(p, q) = \int_0^x t^{p-1}(1-t)^{q-1}\,dt \Big/ \int_0^1 t^{p-1}(1-t)^{q-1}\,dt.$$

From observations on M and D, regression estimates may be obtained, although the use of $I_x(p, q)$ as an expected value presents formidable difficulties. To avoid these difficulties, various attempts were made to find a transformation that would provide an approximate linear relationship between functions of the dependent and independent variables. To be of any value the approximation had to be satisfactory over almost the entire range of $I_x(p, q)$ and had to be valid for values of p, q which were unknown, and for values of x which varied over an unknown range. A search of the literature revealed that, for reasons already cited, no available single formula would suffice. However, by combining two formulae found in the literature and by making one modification, we arrived at a linear transformation that proved adequate for the regression problem. Although it is less accurate than some approximations for the purpose of evaluating $I_x(p, q)$, it has obvious advantages for use as an expected value in the curve-fitting problem.

2. THE APPROXIMATION

In connexion with a sample size problem, Scheffé & Tukey (1944) give, as an approximate solution to the equation $I_x(p, q) = \alpha$,

$$(p+q-1) = \tfrac{1}{4}\chi_\alpha^2\left(\frac{1+x}{1-x}\right) + \tfrac{1}{2}(q-1), \qquad (1)$$

where χ_α^2 is the $100\alpha\,\%$ point of the χ^2-distribution with $2q$ degrees of freedom. They indicate that for $0{\cdot}005 \leqslant \alpha \leqslant 0{\cdot}10$ and $x \geqslant 0{\cdot}9$, the error in this solution is less than $0{\cdot}1\,\%$. It also does well for larger

† Operated by Union Carbide Corporation for the U.S. Atomic Energy Commission.

values of α but becomes inaccurate as x is decreased. We have found that the approximation can be extended to smaller values of x if the first term is multiplied by

$$\left(\frac{21}{23}+\frac{1}{19x}+\frac{1}{3}e^{-5/(2x)}\right)^{-1}. \tag{2}$$

The following modification of the Scheffé–Tukey approximation has been obtained by Tukey (1958)

$$\chi^2_\alpha = 2q(1-\sqrt{x})+(2p-1)\,(1/\sqrt{x}-\sqrt{x}). \tag{3}$$

Computations over the range of arguments in Table 1 show that this new form is remarkably superior to the original approximation. Unfortunately, it does not permit linearization of the regression problem.

Wilson & Hilferty (1931) give the following approximation to χ^2_α,

$$\chi^2_\alpha = 2q\left(1-\frac{1}{9q}+y_{1-\alpha}\sqrt{\frac{1}{9q}}\right)^3, \tag{4}$$

where $y_{1-\alpha}$ is defined by

$$\int_{-\infty}^{y_{1-\alpha}} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}t^2}dt = \alpha.$$

According to the results of Goldberg & Levine (1946), the Wilson–Hilferty approximation, when compared with the three others, does best over a wider range of arguments and hence was chosen for use in our regression problem.

By combining (4) with (1) as modified by (2), one obtains

$$y_\alpha = A+BX, \tag{5}$$

where

$$A=\frac{9q-1}{3\sqrt{q}},\quad B=-3\sqrt{q}\left[1+\frac{2p-1}{q}\right]^{\frac{1}{2}},$$

$$X=\left[\left(\frac{1-x}{1+x}\right)\left(\frac{21}{23}+\frac{1}{19x}+\frac{1}{3}e^{-5/(2x)}\right)\right]^{\frac{1}{2}}.$$

Thus, for any x, p and q, y_α can be computed readily and the approximation,

$$\hat{I} = I_x(p,q) = \int_{-\infty}^{y_\alpha} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}t^2}dt, \tag{6}$$

obtained from tables of the normal distribution. The advantage of this approximation in the aforementioned regression problem should be obvious. Not only is the problem linear, but the dependent variable plus five is a probit, and the literature abounds with computing aids for probit analysis.

By combining (3) with (4), one obtains

$$y_\alpha = A-3\sqrt{q}\left[(1-x)+\frac{2p-1}{2q}\left(\frac{1-x}{\sqrt{x}}\right)\right]^{\frac{1}{2}}. \tag{7}$$

In Table 1, the approximation obtained from (7) is identified as I^*. For evaluating $I_x(p,q)$, this form is definitely superior to (5), but it is of little use in the regression problem.

In Table 1, three approximations are compared with exact values as obtained from Pearson's (1934) tables and from the Harvard tables (1955). The range covered is $2 \leqslant p \leqslant 500$ and $1 \leqslant q \leqslant 451$ for each of four values of x chosen so that no interpolation would be required to obtain the exact values and so that $I_x(p,q)$ would have values near 0·05, 0·30, 0·70 and 0·95. To conserve space, the values of x are not shown. The ranges of p and q correspond to sample sizes between 2 and 951 for the related binomial distribution. Since, for large values of p and q, the direct normal approximation to $I_x(p,q)$ would be expected to do quite well, Table 1 also contains the approximation obtained by taking

$$y_\alpha = \frac{p(1-x)-x(q-1)-0{\cdot}5}{[(p+q-1)\,x(1-x)]^{\frac{1}{2}}}. \tag{8}$$

In Table 1, the approximation obtained from (8) is identified as I'.

3. Goodness of the approximations

Inspection of Table 1 shows that $\hat{I}$ does generally better than I' for values of $p \leqslant 80$ and $q \leqslant 41$ except where p and q are about equal. This latter case corresponds to a probability near one-half for the binomial distribution in which case the normal approximation provides good accuracy even for relatively small

Table 1†. *Approximations to* $I_x(p, q)$

I = exact value
$\hat{I}$ = approximation from (5) and (6)
I' = approximation from (6) and (8)
I^* = approximation from (6) and (7)

(Numbers in parentheses are errors to the nearest *third* decimal.)

p	q		x_1	x_2	x_3	x_4
2	1	I	0·04840	0·30250	0·70560	0·94090
		$\hat{I}$	·05109 (+ 3)	·30871 (+ 6)	·71294 (+ 7)	·93178 (− 9)
		I'	·03519 (−13)	·28483 (−18)	·63577 (−70)	·96591 (+25)
		I^*	·04724 (− 1)	·31094 (+ 8)	·71291 (+ 7)	·93149 (− 9)
2	2	I	0·05331	0·29549	0·70451	0·94669
		$\hat{I}$	·04478 (− 8)	·28806 (− 7)	·70235 (− 2)	·94316 (− 4)
		I'	·03617 (−17)	·30672 (+11)	·69328 (−11)	·96383 (+17)
		I^*	·05120 (− 2)	·30762 (+12)	·71137 (+ 7)	·94360 (− 3)
5	2	I	0·05097	0·30065	0·70441	0·95408
		$\hat{I}$	·05247 (+ 2)	·30372 (+ 3)	·70941 (+ 5)	·95176 (− 2)
		I'	·05074 (0)	·28728 (−13)	·67231 (−32)	·97499 (+21)
		I^*	·04987 (− 1)	·30253 (+ 2)	·70829 (+ 4)	·95134 (− 3)
5	4	I	0·05054	0·29985	0·70640	0·95244
		$\hat{I}$	·04715 (− 3)	·29179 (− 8)	·70221 (− 4)	·95055 (− 2)
		I'	·04470 (− 6)	·30007 (0)	·69808 (− 8)	·96282 (+10)
		I^*	·04887 (− 2)	·30346 (+ 4)	·70939 (+ 3)	·95113 (− 1)
12	2	I	0·04731	0·29205	0·72063	0·94363
		$\hat{I}$	·04884 (+ 2)	·29602 (+ 4)	·72649 (+ 6)	·94207 (− 2)
		I'	·05688 (+10)	·27212 (−20)	·68092 (−40)	·96444 (+21)
		I^*	·04661 (0)	·29187 (0)	·72385 (+ 3)	·94142 (− 2)
12	6	I	0·05167	0·29231	0·70109	0·94521
		$\hat{I}$	·05197 (0)	·29274 (0)	·70258 (+ 2)	·94505 (0)
		I'	·05246 (+ 1)	·28709 (− 5)	·69045 (−11)	·95403 (+ 9)
		I^*	·05093 (− 1)	·29312 (+ 1)	·70265 (+ 2)	·94469 (0)
12	10	I	0·04811	0·29888	0·69144	0·94519
		$\hat{I}$	·04280 (− 5)	·28367 (−15)	·67950 (−12)	·94193 (− 3)
		I'	·04591 (− 2)	·29868 (0)	·68792 (− 4)	·94949 (+ 4)
		I^*	·04701 (− 1)	·30057 (+ 2)	·69304 (+ 2)	·94450 (− 1)
30	5	I	0·04909	0·29168	0·68274	0·94936
		$\hat{I}$	·05149 (+ 2)	·29782 (+ 6)	·68925 (+ 6)	·95013 (+ 1)
		I'	·05657 (+ 8)	·27948 (−12)	·66150 (−21)	·96217 (+13)
		I^*	·04878 (0)	·29138 (0)	·68408 (+ 1)	·94894 (0)
30	15	I	0·05268	0·29271	0·72342	0·95542
		$\hat{I}$	·05389 (+ 1)	·29295 (0)	·72409 (+ 1)	·95540 (0)
		I'	·05413 (+ 1)	·28917 (− 4)	·71840 (− 5)	·96030 (+ 5)
		I^*	·05227 (0)	·29312 (0)	·72409 (+ 1)	·95518 (0)
30	25	I	0·04282	0·29726	0·69384	0·94300
		$\hat{I}$	·03596 (− 7)	·27380 (−24)	·67286 (−21)	·93690 (− 6)
		I'	·04216 (− 1)	·29683 (0)	·69210 (− 2)	·94495 (+ 2)
		I^*	·04215 (− 1)	·29810 (+ 1)	·69469 (+ 1)	·94255 (0)

Table 1 (*cont.*)

p	q		x_1	x_2	x_3	x_4
80	11	I	0·05292	0·26911	0·71251	0·95032
		$\hat{I}$	·05644 (+ 4)	·27806 (+ 9)	·72087 (+ 8)	·95220 (+ 2)
		I'	·05892 (+ 6)	·26175 (− 7)	·70092 (−12)	·95864 (+ 8)
		I^*	·05277 (0)	·26882 (0)	·71307 (+ 1)	·95021 (0)
80	41	I	0·05198	0·30621	0·66294	0·94966
		$\hat{I}$	·05184 (0)	·30574 (0)	·66270 (0)	·94957 (0)
		I'	·05318 (+ 1)	·30380 (− 2)	·69545 (+32)	·95210 (+ 2)
		I^*	·05176 (0)	·30644 (0)	·66336 (0)	·94949 (0)
80	71	I	0·04292	0·31229	0·68943	0·95914
		$\hat{I}$	·03102 (−12)	·26657 (−46)	·64480 (−45)	·94845 (−11)
		I'	·04270 (0)	·31207 (0)	·68877 (− 1)	·95994 (+ 1)
		I^*	·04250 (0)	·31286 (+ 1)	·69000 (+ 1)	·95881 (0)
200	41	I	0·05543	0·33028	0·65056	0·95911
		$\hat{I}$	·06088 (+ 6)	·34758 (+17)	·66684 (+16)	·96240 (+ 3)
		I'	·05833 (+ 3)	·32504 (− 5)	·64422 (− 6)	·96287 (+ 4)
		I^*	·05536 (0)	·33029 (0)	·65079 (0)	·95909 (0)
200	121	I	0·05711	0·31223	0·73254	0·96066
		$\hat{I}$	·05237 (− 5)	·29822 (−14)	·72022 (−12)	·95763 (− 3)
		I'	·05770 (+ 1)	·31107 (− 1)	·73147 (− 1)	·96175 (+ 1)
		I^*	·05694 (0)	·31254 (0)	·73274 (0)	·96051 (0)
200	181	I	0·03961	0·27940	0·72157	0·96143
		$\hat{I}$	·02318 (−16)	·20972 (−70)	·64741 (−74)	·94193 (−20)
		I'	·03956 (0)	·27930 (0)	·72129 (0)	·96179 (0)
		I^*	·03935 (0)	·27969 (0)	·72184 (0)	·96120 (0)
500	101	I	0·02152	0·21393	0·69515	0·97166
		$\hat{I}$	·02607 (+ 5)	·23660 (+23)	·71934 (+24)	·97562 (+ 4)
		I'	·02328 (+ 2)	·21273 (− 1)	·69186 (− 3)	·97389 (+ 2)
		I^*	·02150 (0)	·21391 (0)	·69533 (0)	·97162 (0)
500	301	I	0·02364	0·20249	0·63018	0·93511
		$\hat{I}$	·02029 (− 3)	·18566 (−17)	·60799 (−22)	·92779 (− 7)
		I'	·02403 (0)	·20225 (0)	·62912 (− 1)	·93569 (+ 1)
		I^*	·02357 (0)	·20236 (0)	·63030 (0)	·93507 (0)
500	451	I	0·05592	0·36063	0·60280	0·93303
		$\hat{I}$	·02576 (−30)	·24218 (−118)	·47013 (−133)	·88040 (−53)
		I'	·05594 (0)	·36048 (0)	·60258 (0)	·93319 (0)
		I^*	·05578 (0)	·36096 (0)	·60313 (0)	·93294 (0)

† Abridged from a larger table available from the author on request.

sample sizes. On the other hand, the approximation given by $\hat{I}$ is not too bad for these cases, particularly for purposes of regression, and its accuracy tends to improve as x increases. As expected, for $p > 80$ and $q > 41$, the normal approximation is uniformly better than $\hat{I}$ and in fact is good enough for almost any practical purpose including the regression problem.

The approximation I^* is surprisingly accurate over the ranges examined and even competes successfully with I' for large p and q. Although $\hat{I}$ does not have the uniform accuracy of I^*, or of I' for large p and q, it is definitely preferable for use in the regression problem in the lower ranges of p and q. Perhaps it should be emphasized that all of these approximations may be used, without the need for interpolation, for integral or fractional values of p and q, the only tables required being the readily available tables of the normal distribution.

REFERENCES

Goldberg, Henry & Levine, Harriet (1946). Approximate formulas for the percentage points and normalization of t and χ^2. *Ann. Math. Statist.* **17**, 216–25.

Harvard Computation Laboratory (1955). *Tables of the Cumulative Binomial Probability Distribution.* Cambridge, Massachusetts: Harvard University Press.

Pearson, Karl (1934). *Tables of the Incomplete Beta Function.* Cambridge, England: The University Press.

Scheffé, H. & Tukey, J. W. (1944). A formula for sample sizes for population tolerance limits. *Ann. Math. Statist.* **15**, 217.

Tukey, J. W. (1958). Private communication.

Wilson, Edwin B. & Hilferty, Margaret M. (1931). The distribution of chi-square. *Proc. Nat. Acad. Sci.* **17**, 684–8.

Wise, M. E. (1950). The incomplete beta function as a contour integral and a quickly converging series for its inverse. *Biometrika*, **37**, 208–18.

Sequential occupancy

By D. E. BARTON and F. N. DAVID

University College London

Introduction

In previous papers (Barton & David, 1959*a*, *b*) we have discussed the distribution of the number of empty compartments when n balls are randomly and independently thrown into N identical compartments. Various restrictions on the compartments have been envisaged. The alternative problem where the number of empty compartments is fixed and the number of balls becomes the random variable is not without interest. The classical case corresponding to De Moivre's differences of zero distribution is, of course, well known, but we think that this is not true of the other cases discussed below.

The actual distributions are simply and easily obtained. Suppose k zero compartments and let the distribution of n, the number of balls required to fill just $N-k$ compartments, be defined. Then the probability of n is equal to the probability that $n-1$ balls will fill $N-k-1$ compartments times the probability that one further ball falls in an empty compartment. This argument will hold whether there is restriction on the compartment size or not, and the sequential distributions thus follow from our previous work (1959*a*). The generalization of Pólya's scheme has also been noted previously.

The moments of the distribution of n are not easy to find unless the actual probability process is investigated (see for example Feller, 1957, p. 210). We discuss here methods of obtaining these moments.

Classical case with no restraint on balls or compartments

The probability distribution of n is

$$p(n) = \frac{1}{N^{n-1}} {}^{N}C_{k+1} \Delta^{N-k-1} 0^{n-1} \frac{k+1}{N} = \frac{1}{N^{n-1}} {}^{N-1}C_{k} \Delta^{N-k-1} 0^{n-1} \quad (n = N-k, N-k+1, \ldots, \infty),$$

a fact which is well known. The moments of n are straightforward. Suppose the dropping of the balls is going on and we break in at the point when the $(N-t)$ th box, empty before, receives one ball leaving $t(\leqslant k)$ boxes empty. Let there be n_i balls in the $(N-t)$ boxes. Let the further number of balls dropped

up to and including the occasion when one further empty box receives one ball be α_{N-t}. Then the total number of balls necessary to achieve $N-k$ filled and k empty is

$$n = \sum_{t=k+1}^{N} \alpha_{N-t},$$

with $\alpha_0 = 1$. α_{N-t} is a sequential binomial variable with the trials stopping when just one success is achieved. Accordingly

$$K_{[r]}(\alpha_{N-t}) = (r-1)!\left(\left(\frac{q}{p}\right)^r + (-1)^{r-1}\right)$$

with

$$p = \frac{t}{N}, \quad q = \frac{N-t}{N},$$

or

$$K_{[r]}(\alpha_{N-t}) = (r-1)!\left(\left(\frac{N-t}{t}\right)^r + (-1)^{r-1}\right).$$

In this classical scheme the α's are mutually independent and consequently

$$K_{[r]}(n) = (r-1)!\left((N-k)(-1)^{r-1} + \sum_{t=k+1}^{N}\left(\frac{N-t}{t}\right)^r\right),$$

or

$$K_{[r]}(n) = (r-1)!\left((N-k)(-1)^{r-1} + \sum_{j=0}^{N-k-1}\left(\frac{j}{N-j}\right)^r\right)$$

(a form which leads more easily to asymptotic expression).

Both the distribution and the factorial cumulants of the classical sequential set-up discussed in the preceding paragraph are well known. It is known also that for k not too large

$$K_{[1]}(n) \doteqdot N\log\frac{N}{k+\frac{1}{2}}, \quad K_{[2]}(n) \doteqdot N^2\left(\frac{\pi^2}{6} - \sum_{r=1}^{k}\frac{1}{r^2}\right).$$

When $k = 0$ the cumulants of

$$x = \frac{n - \log N}{N}$$

tend with increasing N to

$$K_1 = \gamma(\text{Euler's constant}), \quad K_r = \sum_{j=1}^{\infty}\frac{1}{j^r} \quad (r \geqslant 2).$$

This implies that in the limit

$$K_r = \left(-\frac{\partial}{\partial z}\right)^r \log\Gamma(z)\Big|_{z=1}.$$

These cumulants are the cumulants of a variable x whose cumulative probability distribution function is

$$\exp(-\exp(-x))$$

and this therefore is the limiting C.P.D.F.

The normal limit follows in the sequential case as in the fixed sample size, classical type occupancy problem.

Let $k = N(1-p)$ and allow N to increase without limit. The mth factorial cumulant of n can be written

$$K_{[m]} = (-1)^{m-1}(N-k)(m-1)! + \sum_{V=0}^{N-k-1}\frac{(V/N)^m(m-1)!}{(1-V/N)^m}$$

$$= (m-1)!\sum_{j=1}^{m} {}^mC_j(-1)^{m+j}\sum_{V=0}^{N-k-1}\left(1-\frac{V}{N}\right)^j$$

Let

$$T_j = \sum_{V=0}^{N-k-1}\left(1-\frac{V}{N}\right)^{-j},$$

$$\frac{T_1}{N} = \sum_{V=0}^{N-k-1}\left(\frac{1}{N-V}\right) \to \log\frac{N}{k} = \log\frac{1}{q}$$

if $q = 1-p$.

$$\frac{T_j}{N} \to \int_0^p \frac{dx}{(1-x)^j} = -\frac{1}{j-1}\left(1-\frac{1}{q^{j-1}}\right) \quad (\text{for } j \geqslant 2).$$

It follows that

$$K_{[m]} \doteqdot (m-1)!\,N\left\{\left(\log\frac{1}{q}\right)m(-1)^{m-1}+\sum_{j=2}^{m}\frac{{}^{m}C_{j}(-1)^{m+j+1}}{j-1}\left(1-\frac{1}{q^{j-1}}\right)\right\}$$

and

$$K_1 \doteqdot N\log\frac{1}{q}$$

$$K_2 \doteqdot -N\left(\log\frac{1}{q}\right)+\frac{pN}{q}.$$

As in the classical type problem this implies that

$$\frac{n-N\log(1/q)}{\sqrt{N\{(p/q)-\log(1/q)\}}}$$

is, asymptotically, a unit normal variable as $N\to\infty$.

Romanovsky's case with restricted compartment capacity

In this case, as we have described in our previous paper (Barton & David, 1959*a*), it is supposed that the capacity of each compartment is restricted to r balls. The probability distribution of n is, for k compartments, to be left empty,

$$\begin{aligned}p(n) &= \frac{r(k+1)}{Nr-n+1}\,\frac{{}^{N}C_{k+1}}{{}^{Nr}C_{n-1}}\sum_{j=0}^{N-k-1-\{(n-1)/r\}}(-1)^{j}\,{}^{N-k-1}C_{j}\,{}^{r(N-k-1-j)}C_{n-1}\\ &= \frac{{}^{N-1}C_{k}}{{}^{Nr-1}C_{n-1}}\sum_{j=0}^{N-k-1-\{(n-1)/r\}}(-1)^{j}\,{}^{N-k-1}C_{j}\,{}^{r(N-k-1-j)}C_{n-1},\end{aligned}$$

a distribution which we believe is new. For the moments we proceed on similar, but not identical, lines to those for the classical sequential case so far considered. Suppose we enter the process at the point where t of the cells have been occupied, by S_t balls, and that n_t balls have to be dropped between the occupation of the tth and the $(t+1)$st cells. We have, conditional on S_t (with conventionally $S_0=0$), that

$$p(n_t)=\frac{(rt-S_t)^{(n_t-1)}}{(Nr-S_t)^{(n_t)}}\,r(N-t)=-\frac{(rt-S_t)^{(n_t)}}{(Nr-S_t)^{(n_t)}}+\frac{(rt-S_t)^{(n_t-1)}}{(Nr-S_t)^{(n_t-1)}}.$$

It follows, still conditional on S_t, that

$$\mathscr{E}\left(\frac{(Nr-S_t-n_t+1)^{[m]}}{(Nr-S_t+1)^{[m]}}\right)=\frac{r(N-t)}{r(N-t)+m},$$

where

$$A^{[m]}=(A+m-1)^{(m)}$$

for any number A and integer m. For the overall expectation we have

$$\mathscr{E}\,(Nr-S_{t+1}+1)^{[m]}=\frac{N^{(t+1)}r^{t+1}(Nr+1)^{[m]}}{P_{m,t+1}},$$

where

$$P_{m,t+1}=\prod_{j=0}^{t}(m+r(N-j)).$$

Reduction of this expectation gives us for the factorial advancing moment of S_t,

$$\mathscr{E}(S_t^{[m]})=(Nr+1)^{[m]}N^{(t)}r^{t}\sum_{i=0}^{m}\frac{{}^{m}C_{i}(-1)^{i}}{P_{i,t}}.$$

n, the number of balls required to achieve just $N-k$ occupied boxes, will be equal to S_{N-k}. We have

$$\mathscr{E}(S_{N-k})=\mathscr{E}(n)=(Nr+1)\left[1-\frac{N!}{\{N+(1/r)\}!}\,\frac{\{k+(1/r)\}!}{k!}\right]$$

and

$$\mathscr{E}(n(n+1))=(Nr+1)(Nr+2)\left[1-\frac{2N!}{\{N+(1/r)\}!}\,\frac{\{k+(1/r)\}!}{k!}+\frac{N!}{\{N+(2/r)\}!}\,\frac{\{k+(2/r)\}!}{k!}\right],$$

where we use the notation $x!=\Gamma(x+1)$ for any $x>0$. The higher factorial moments follow similarly.

Apparent contagion within a compartment; a generalized Pólya scheme

Following our previous paper (1959*a*) if a is a measure of the contagion within a cell then the probability that $(k-1)$ cells out of N will be left empty when $(n-1)$ balls are dropped randomly is

$$\frac{{}^{N}C_{k+1}}{{}^{aN+n-2}C_{n-1}} \sum_{j=0}^{N-k-1} (-1)^j {}^{N-k-1}C_j \, {}^{a(N-k-1-j)+n-2}C_{n-1}.$$

As we remarked there, this is a generalization of Pólya's scheme. The probability that one further ball drops in an empty compartment is

$$\frac{a(k+1)}{aN+n-1},$$

so that the probability distribution of n, the number of balls required to occupy $(N-k)$ compartments, is

$$p(n) = \frac{{}^{N-1}C_{k}}{{}^{aN+n-1}C_{n-1}} \sum_{j=0}^{N-k-1} (-1)^j {}^{N-k-1}C_j \, {}^{a(N-k-1-j)+n-2}C_{n-1}.$$

The moments follow by arguments similar to those which we have used for the case immediately preceding. Alternatively, we note that formally we may write

$$a = -r, \quad (-A)^{(m)} = (-1)^m A^{[m]}.$$

Thus
$$\mathscr{E}\left(\frac{(Na+S_t+n_t-1)^{(m)}}{(Na+S_t-1)^{(m)}}\right) = \frac{(N-t)a}{a(N-t)-m}$$

conditional on S_t. The overall expectation is

$$\mathscr{E}(Na+S_{t+1}-1)^{(m)} = \frac{N^{(t+1)}a^{t+1}(Na-1)^{(m)}}{P_{m,t+1}},$$

where
$$P_{m,t+1} = \prod_{j=0}^{t} (a(N-j)-m).$$

Generally
$$\mathscr{E}(S_t^{(m)}) = (Na-1)^{(m)} N^{(t)} a^t \sum_{i=0}^{m} \frac{{}^{m}C_i (-1)^{i+m}}{P_{i,t}}.$$

Generalized sequential occupancy: unequal probabilities in each box

We consider here a multinomial type of sequential occupancy where there is no restriction as to box size but the probabilities for each box are not the same. Assume as before N boxes with probabilities $p_1, p_2, \ldots, p_N$ of the balls falling in them. The balls are thrown independently until $(N-k)$ boxes have been occupied.

(i) *n fixed and k, the number of zeros, a random variable*

For this, the non-sequential case, we define a set of random variables α such that

$\alpha_i = 1$ if the ith box is empty after n balls are dropped,
$= 0$ if not.

Then the number of empty boxes is

$$k = \sum_{i=1}^{N} \alpha_i$$

and
$$\mathscr{E}(k^{(m)}) = m! \sum_{i_1<i_2<\ldots<i_m} (1-p_{i_1}-\ldots-p_{i_m})^n$$
$$= m! S_m^{(n)} (\text{say}) \quad (m \geqslant 1)$$

with
$$S_N^{(n)} = 0.$$

The probability generating function of k is

$$\Pi(z) = 1 + \sum_{m=1}^{N-1} (z-1)^m S_m^{(n)}$$

(ii) *k fixed and n a random variable*

Let $P_k(n)$ denote the cumulative probability distribution function of n. We have

$$\begin{aligned} P_k(n) &= \text{probability of at most } k \text{ zeros after } n \text{ throws} \\ &= \text{coefficient of } z^k \text{ in } \pi(z)/(1-z). \end{aligned}$$

Write $$\pi^*(y) = \Sigma p_k(n)\, y^n,$$

where $p_k(n)$ is the P.D.F. of n. We have

$$\begin{aligned} \Sigma P_k(n)\, y^n &= \frac{\pi^*(y)}{1-y} \\ &= \text{coefficient of } z^k \text{ in } \frac{1}{1-z}\left\{\frac{1}{1-y} + \sum_{m=1}^{N-1} (z-1)^m R_m(y)\right\}, \end{aligned}$$

where $$R_m(y) = \sum_{i_1<i_2<\ldots<i_m} \frac{1}{1-(1-p_{i_1}-\ldots-p_{i_m})y}.$$

It follows that $$\frac{\pi^*(y)-1}{y-1} = \text{coefficient of } z^k \text{ in } \sum_{m=1}^{N-1} (z-1)^{m-1} R_m(y).$$

If further we put $y = 1+\theta$, $$\frac{\pi^*(y)-1}{y-1} = \sum_{V=1}^{\infty} \frac{\mu_{[V]}}{V!}\theta^{V-1}$$

and therefore

$$\frac{\mu_{[V]}}{V!} = \text{coefficient of } z^k \text{ in } \sum_{m=1}^{N-1} (z-1)^{m-1} \sum_{i_1<\ldots<i_m} \frac{[1-(p_{i_1}+\ldots+p_{i_m})]^{V-1}}{(p_{i_1}+\ldots+p_{i_m})^V}.$$

TWO APPLICATIONS OF THE PREVIOUS SECTION

(i) Assume a line of unit length is broken, divided randomly into N pieces of lengths $p_1, p_2, \ldots, p_N$. If y_m is the sum of m of these, then

$$p(y_m) = \frac{(N-1)!}{(m-1)!\,(N-m-1)!}\, y^{m-1}(1-y)^{N-m-1}.$$

If these lengths $p_1, p_2, \ldots, p_N$ are now the box probabilities then the Vth factorial moment of n, the number of balls required to occupy $N-k$ boxes is given by $\mu_{[V]}/V!$ which is the coefficient of z^k in

$$\sum_m (z-1)^{m-1}\, {}^N C_m \frac{(N-m)^{[V]}}{(m-1)^{(V-1)}}(N-1).$$

(ii) Let the box probabilities bear a linear relationship one to the other so that

$$p_j = 2j/(N(N+1)) \quad (j = 1, 2, \ldots, N).$$

Let us consider the mean number of balls necessary to ensure that every box is occupied. We have generally

$$\mu_{[1]} = \sum_m (-1)^{m-1} \Sigma \frac{1}{p_{i_1}+\ldots+p_{i_m}},$$

where the inner sum is over all possible sets of m different p's. For our particular case

$$\begin{aligned} ({}^{N+1}C_2)^{-1}\mu_{[1]} &= \sum_{m=1}^{N} (-1)^{m-1} \Sigma \frac{1}{i_1+\ldots+i_m} \\ &= \int_0^1 \frac{\prod_{i=1}^{N}(1-x^i)-1}{x}\, dx \end{aligned}$$

which may be shown to tend to $\{(4\pi/\sqrt{3})-6\}$. Thus

$$\mu_{[1]} \sim {}^{N+1}C_2\left(\frac{4\pi}{\sqrt{3}} - 6\right).$$

This mean value is very much larger than that for the classical sequential occupancy scheme, where if

$$p_i = \frac{1}{N}$$

we have

$$\mu_{[1]} \sim N \log N.$$

Ever since the early paper of Stevens (1937) the distribution of the number of zeros in the classical (De Moivre) case has been used in practice. Recently the Romanovsky type of problem has come to the front as, for example, where Turner & Eadie (1957) discussed the distribution of cells in haemacytometer counts when no overlapping of cells is permitted. So far we have not seen any applications of the sequential occupancy distributions other than in slightly unreal problems such as those connected with cigarette cards. There is no doubt, however, that sequential distributions are implicit in any collectors' problem. For example, the compartments may be traps and the arrival of the insects (i) may be random, (ii) may be conditioned by whether there is room in the traps, or (iii) may be decided by the equivalent of a gregarious spirit among the insects. These three conditions will correspond respectively to the three cases we have discussed, namely, the classical, the Romanovsky and the generalized Pólya scheme. It is also possibly worth noting that the sequential aspect of the occupancy problem may be considered as the number of renewals in a special form of renewal process.

REFERENCES

Barton, D. E. & David, F. N. (1959*a*). *J.R. Statist. Soc.* B. (In the Press.)

Barton, D. E. & David, F. N. (1959*b*). *J.R. Statist. Soc.* B. (In the Press.)

Feller, W. (1957). *An Introduction to Probability Theory.* New York: John Wiley and Sons.

Stevens, W. L. (1937). *Ann. Eugen., Lond.*, **8**, 57–69.

Turner, M. & Eadie, G. (1957). *Biometrics*, **13**, 485–95.

The generalized mean differences of the binomial and Poisson distributions

By T. A. RAMASUBBAN

London School of Economics and Political Science

1. Introduction

Gini's mean difference g, defined by

$$g = \frac{2}{n(n-1)} \sum_i \sum_{j \neq i} |x_i - x_j| \tag{1·1}$$

is one of the well-known and useful measures of dispersion. Unlike most of the quantities designed with a view to measuring the amount of scatter, g is independent of any central measure of location. But, owing to the mathematical complexities introduced by its definition, the study of the distribution of this statistic and in particular some of its sampling properties has not advanced quite as much as in the case of other scale measures, except perhaps when the variates follow a normal law.

For the normal population, the exact standard error of g had been first given by Nair (1936) by a rather complicated method. Later Lomnicki (1952) obtained the same by a simpler derivation. Kamat (1953) evaluated the third moment in an exact form and, from the considerations of the measure of skewness β_1, conjectured that for large n the distribution of g may be that of a χ. Following Kamat, I had obtained (1956) approximate values for its fourth moment and showed that the measure of kurtosis β_2 calculated therefrom, taken together with the values of β_1 obtained by Kamat, seem to establish the closeness of the χ-approximation, at least for sample size $n > 10$. I have also attempted to obtain empirical distributions of g for small samples ($n < 10$), the results of which I propose to report at a later date.

The normal distribution can be looked upon as a limiting case of the binomial and Poisson distributions under certain conditions. It is therefore natural to investigate the sampling moments of g for these discrete distributions and to attempt fitting an appropriate distribution for g even if the exact distribution is not easily found. In order to facilitate this work, the first step is to derive expressions for the absolute mean differences Δ_r given by

$$\Delta_r = \sum_i \sum_{j \neq i} |i-j|^r p(i)\, p(j) \tag{1·2}$$

for a few of the values of r, where i and j take all the integral values permitted by the discrete distribution $p(t)$. In a recent article (1958), I have succeeded in obtaining values of Δ_1 for some of the discontinuous distributions; in particular, those for the binomial and the Poisson distributions are shown to be:

$$\Delta_1(\text{binomial}) = 2npq\left[\sum_{i=0}^{n-1}\binom{n-1}{i}^2 p^{2i}q^{2n-2i-2} + \sum_{i=1}^{n-1}\binom{n-1}{i}\binom{n-1}{i-1}p^{2i-1}q^{2n-2i-1}\right] \tag{1·3}$$

and

$$\Delta_1(\text{Poisson}) = 2\lambda e^{-2\lambda}[I_0(2\lambda)+I_1(2\lambda)], \tag{1·4}$$

where $I_n(x)$ is the nth order modified Bessel function of the first kind. In the present paper I shall evaluate for these two distributions the generalized expressions for Δ_r valid for integral values of r.

2. The even order mean differences

By definition, the even order mean difference is given by

$$\Delta_{2r} = \sum_i\sum_j |i-j|^{2r}p(i)\,p(j) = \sum_i\sum_j (i-j)^{2r}p(i)\,p(j). \tag{2·1}$$

Let ξ be the arithmetic mean and μ_k be the kth moment about ξ. Then, since $\mu_1 = 0$,

$$\begin{aligned}\Delta_{2r} &= \sum_i\sum_j\{(i-\xi)-(j-\xi)\}^{2r}p(i)\,p(j) = \mu_{2r}+\binom{2r}{2}\mu_{2r-2}\mu_2-\binom{2r}{3}\mu_{2r-3}\mu_3+\ldots+\binom{2r}{2r-2}\mu_2\mu_{2r-2}+\mu_{2r}\\ &= 2\mu_{2r}+\sum_{j=2}^{2r-2}(-1)^j\binom{2r}{j}\mu_{2r-j}\mu_j.\end{aligned} \tag{2·2}$$

It is thus possible to express the mean difference of even order in terms of the moments of the distribution. In particular,

$$\left.\begin{aligned}\Delta_2 &= 2\mu_2,\\ \Delta_4 &= 2(\mu_4+3\mu_2^2),\\ \Delta_6 &= 2(\mu_6+15\mu_4\mu_2-10\mu_3^2),\quad\text{and so on}\end{aligned}\right\} \tag{2·3}$$

On substituting the various values for the moments of the binomial and the Poisson distributions in (2·2) or (2·3), the corresponding even order mean differences can be evaluated.

3. The odd order mean differences for the binomial distribution

The evaluation of the odd order mean differences for the binomial distribution calls for some interesting algebraic manipulations. We have

$$\Delta_{2r+1} = \sum_{i=0}^{n}\sum_{j=0}^{n}|i-j|^{2r+1}\binom{n}{i}\binom{n}{j}p^iq^{n-i}p^jq^{n-j} \tag{3·1}$$

and since the sum of $(i-j)^{2r+1}$ over the distribution vanishes, (3·1) becomes

$$\Delta_{2r+1} = 2\sum_{i=0}^{n}\sum_{j=0}^{i}(i-j)^{2r+1}\binom{n}{i}\binom{n}{j}p^iq^{n-i}p^jq^{n-j} \tag{3·2}$$

$$= 2\sum_{i=0}^{n}\binom{n}{i}p^iq^{n-i}\left\{\sum_{j=0}^{i}(i-j)^{2r+1}\binom{n}{j}p^jq^{n-j}\right\}$$

$$= 2\sum_{i=0}^{n}\binom{n}{i}p^iq^{n-i}\left\{i^{2r+1}\binom{n}{0}q^n+(i-1)^{2r+1}\binom{n}{1}pq^{n-1}+\ldots+\binom{n}{i-1}p^{i-1}q^{n-i+1}\right\}. \tag{3·3}$$

Expanding the sum in (3·3) and rearranging the terms to form a finite series in p^sq^{2n-s} $(s = 1, 2, \ldots, 2n-1)$ we have

$$\begin{aligned}\Delta_{2r+1} = 2\Bigg[\sum_{i=0}^{n-1}\left\{\sum_{j=0}^{i}(2j+1)^{2r+1}\binom{n}{i-j}\binom{n}{i+j+1}\right\}p^{2i+1}q^{2n-2i-1}\\ +\sum_{i=1}^{n-1}\left\{\sum_{j=1}^{i}(2j)^{2r+1}\binom{n}{i-j}\binom{n}{i+j}\right\}p^{2i}q^{2n-2i}\Bigg].\end{aligned} \tag{3·4}$$

The method of evaluation of the sums

$$\sum_{j=0}^{i}(2j+1)^{2r+1}\binom{n}{i-j}\binom{n}{i+j+1} \quad\text{and}\quad \sum_{j=1}^{i}(2j)^{2r+1}\binom{n}{i-j}\binom{n}{i+j}$$

is somewhat involved and lengthy. I shall therefore omit such of the algebraical details which may not impair its clarity.

Consider the sum $\sum_{j=1}^{i} (2j)^{2r+1} \binom{n}{i-j}\binom{n}{i+j}$.

Let
$$S_k = \sum_{j=1}^{i-k} \binom{n-k}{i-j-k}\binom{n-k}{i+j} t^{2j+k}$$

and
$$\sigma_{k,l} = \sum_{j=1}^{i-k} (2j+k)^l \binom{n-k}{i-j-k}\binom{n-k}{i+j} t^{2j+k} = D^l(S_k),$$

where the operator $D \equiv t(\partial/\partial t)$. The sum required is therefore $\sigma_{0,2r+1}|_{t=1} = D^{2r+1}S_0|_{t=1}$; and since the expanded form of $D^{2r+1}S_0$ has $DS_0 (= \sigma_{0,1})$ as a factor, it is more convenient to operate $2r$ times on $\sigma_{0,1}$ and then set $t = 1$. Now

$$\sigma_{0,1} = \sum_{j=1}^{i} (2j) \binom{n}{i-j}\binom{n}{i+j} t^{2j}. \tag{3·5}$$

Splitting $2j$ as $\{(i+j)-(i-j)\}$, using Vandermonde's theorem, and finally summing over j, (3·5) gives

$$\sigma_{0,1} = n\left[t^2\binom{n-1}{i}\binom{n-1}{i-1} + \left(t-\frac{1}{t}\right)\sum_{j=1}^{i-1}\binom{n-1}{i-j-1}\binom{n-1}{i+j} t^{2j+1}\right]. \tag{3·6}$$

On putting $t = 1$ in (3·6), we have the well-known result

$$\sigma_{0,1}|_{t=1} = n\binom{n-1}{i}\binom{n-1}{i-1}.$$

Again,

$$\sigma_{0,2r+1}|_{t=1} = D^{2r}\sigma_{0,1}|_{t=1} = n\left[\binom{n-1}{i}\binom{n-1}{i-1} D^{2r}(t^2) + \sum_{j=1}^{i-1}\binom{n-1}{i-j-1}\binom{n-1}{i+j} D^{2r}\left\{\left(t-\frac{1}{t}\right)t^{2j+1}\right\}\right]_{t=1}. \tag{3·7}$$

It can readily be seen that any even number of operations by D on $(t-t^{-1})$ vanishes after t is put equal to unity and that any odd number of operations yields a factor 2. Therefore, expanding the term $D^{2r}\{(t-t^{-1})t^{2j+1}\}$ on the extreme right of (3·7), by Leibniz's theorem and letting $t = 1$, we have

$$\begin{aligned} D^{2r}\{(t-t^{-1})t^{2j+1}\}|_{t=1} &= 2\left[\left\{\binom{2r}{1}D^{2r-1} + \binom{2r}{3}D^{2r-3} + \ldots + \binom{2r}{r-1}D^{r+1}\right\}t^{2j+1}\right]_{t=1} \\ &= 2\sum_{p_1=1}^{r}\binom{2r}{2p_1-1}(2j+1)^{2r-2p_1+1} \end{aligned} \tag{3·8}$$

(3·7) can therefore be written as

$$\sigma_{0,2r+1}\Big|_{t=1} = n\binom{n-1}{i}\binom{n-1}{i-1}2^{2r} + 2n\sum_{p_1=1}^{r}\binom{2r}{2p_1-1}D^{2r-2p_1+1}S_1\Big|_{t=1}. \tag{3·9}$$

Now
$$DS_1 = \sum_{j=1}^{i-1}(2j+1)\binom{n-1}{i-j-1}\binom{n-1}{i+j}t^{2j+1}.$$

Splitting $(2j+1)$ as $\{(i+j)-(i-j-1)\}$ and summing as above we find,

$$DS_1 = (n-1)\left[\binom{n-2}{i}\binom{n-2}{i-2}t^3 + \left(t-\frac{1}{t}\right)\sum_{j=1}^{i-2}\binom{n-2}{i-j-2}\binom{n-2}{i+j}t^{2j+2}\right]. \tag{3·10}$$

Operating on (3·10), $(2r-2p_1)$ times exactly as before and letting $t = 1$, we may show that

$$D^{2r-2p_1+1}S_1 = (n-1)\binom{n-2}{i}\binom{n-2}{i-2}3^{2r-2p_1} + 2(n-1)\sum_{p_2=1}^{r-p_1}\binom{2r-2p_1}{2p_2-1}D^{2r-2p_1-2p_2+1}S_2\Big|_{t=1}. \tag{3·11}$$

Substitution of (3·11) in (3·9) and simplification yields,

$$\begin{aligned} \sigma_{0,2r+1}\Big|_{t=1} = {} & a_{r1}n^{(1)}\binom{n-1}{i}\binom{n-1}{i-1} + a_{r2}n^{(2)}\binom{n-2}{i}\binom{n-2}{i-2} \\ & + 2^2.n^{(2)}\sum_{p_1=1}^{r}\sum_{p_2=1}^{r-p_1}\binom{2r}{2p_1-1}\binom{2r-2p_1}{2p_2-1}D^{2r-2p_1-2p_2+1}S_2\Big|_{t=1}, \end{aligned} \tag{3·12}$$

where $$a_{r1} = 2^{2r} \quad \text{and} \quad a_{r2} = \frac{2^{2r+1}}{3.4}(2^{2r+1}-2)$$

and $n^{(s)} = n(n-1)(n-2)\dots(n-s+1)$.

Dealing with (3·12) in exactly the same manner as with (3·9) whose forms are identical, we finally have

$$\sigma_{0,2r+1}\Big|_{t=1} = \sum_{s=1}^{r+1} a_{rs} n^{(s)} \binom{n-s}{i}\binom{n-s}{i-s}, \tag{3·13}$$

where $$a_{rs} = 2^{s-1} \sum_{p_1=1}^{r} \sum_{p_2=1}^{r-p_1} \cdots \sum_{p_{s-1}=1}^{r-\sum_{h=1}^{s-2} p_h} \binom{2r}{2p_1-1}\binom{2r-2p_1}{2p_2-1}\cdots\binom{2r-\sum_{h=1}^{s-2} 2p_h}{2p_{s-1}-1}(s+1)^{2r-\sum_{h=1}^{s-1} 2p_h}. \tag{3·14}$$

It will be evident from (3·14) that the series (3·13) must terminate at $s = r+1$ since, for values of s greater than this, the last binomial coefficient occurring in (3·14) will vanish, even for the minimum value of the p_h, namely unity.

The evaluation of a_{rs} in (3·14) is not as difficult as it appears to be. The successive summations, starting with respect to p_{s-1}, ultimately give

$$a_{rs} = \frac{2^{2r+1}}{(s+1)\dots(2s)} \sum_{m=0}^{s-1} (-1)^m \left\{\binom{2s-1}{m} - \binom{2s-1}{m-1}\right\}(s-m)^{2r+1}. \tag{3·15}$$

The procedure being the same for effecting the other sum, $\sum_{j=0}^{i} (2j+1)^{2r+1}\binom{n}{i-j}\binom{n}{i+j+1}$, I shall simply state the final result:

$$\sum_{j=0}^{i} (2j+1)^{2r+1}\binom{n}{i-j}\binom{n}{i+j+1} = \sum_{s=1}^{r+1} a'_{rs} n^{(s)} \binom{n-s}{i}\binom{n-s}{i-s+1}, \tag{3·16}$$

where $$a'_{r,s+1} = \frac{1}{(s+1)\dots(2s)} \sum_{m=0}^{s} (-1)^m \left\{\binom{2s}{m} - \binom{2s}{m-1}\right\}(2s-2m+1)^{2r}. \tag{3·17}$$

It may be pointed out here that there appears to be more than one method of evaluating these sums and that each method provides series similar to (3·13) and (3·16). However, these different series are only variations of the ones given here with the coefficients a_{rs} and a'_{rs} modified. It also seems possible to obtain a few relations concerning these coefficients but they are of only algebraic interest. The only relation which I shall make use of at a later stage is

$$a_{r,r+1} = a'_{r,r+1} = 2^{2r} r! \tag{3·18}$$

On substituting (3·13) and (3·16) in (3·4), the mean difference becomes

$$\Delta_{2r+1} = 2\left[\sum_{i=0}^{n-1}\left\{\sum_{s=1}^{r+1} a'_{rs} n^{(s)}\binom{n-s}{i}\binom{n-s}{i-s+1}\right\}p^{2i+1}q^{2n-2i-1} + \sum_{i=1}^{n-1}\left\{\sum_{s=1}^{r+1} a_{rs} n^{(s)}\binom{n-s}{i}\binom{n-s}{i-s}\right\}p^{2i}q^{2n-2i}\right]. \tag{3·19}$$

It can be verified that on putting $r = 0$ in (3·19), the value of Δ_1 given by (1·3) is readily obtained.

(3·19) can be modified to give an alternative form for Δ_{2r+1} in terms of the hypergeometric function ${}_2F_1(\alpha, \beta; \gamma; x)$, where

$$F(\alpha, \beta; \gamma; x) = 1 + \frac{\alpha\beta}{\gamma}x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)}\frac{x^2}{2!} + \dots.$$

Taking account of the fact that the limits for i in (3·19) are to be taken such that none of the binomial coefficients like $\binom{n-s}{s}$ vanish, we find

$$\Delta_{2r+1} = 2\sum_{s=1}^{r+1}\left\{a'_{rs} n^{(s)} p^s q^s \sum_{i=s-1}^{n-s}\binom{n-s}{i}\binom{n-s}{i-s+1}p^{2i-s+1}q^{2n-2i-s-1} + a_{rs} n^{(s)} p^s q^s \sum_{i=s}^{n-s}\binom{n-s}{i}\binom{n-s}{i-s}p^{2i-s}q^{2n-2i-s}\right\}. \tag{3·20}$$

It is then not difficult to show that the sums in the braces on the right of (3·20) are respectively equal to

$$\binom{n-s}{s-1}(pq)^{s-1}\,{}_2F_1(-n+2s-1, \tfrac{1}{2}(2s-1); 2s-1; 4pq)$$

and

$$\binom{n-s}{s}(pq)^{s}\,{}_2F_1(-n+2s, \tfrac{1}{2}(2s+1); 2s+1; 4pq),$$

from the fact that these sums are the coefficients of $1/t^{s-1}$ and $1/t^s$ in $\{(q+pt)(q+p/t)\}^{n-s}$, i.e. in

$$\left\{1+\frac{(1-t)^2}{t}pq\right\}^{n-s}, \quad \text{since } q+p=1.$$

Thus,

$$\Delta_{2r+1} = 2\sum_{s=1}^{r+1}\left\{a'_{rs}n^{(s)}\binom{n-s}{s-1}p^{2s-1}q^{2s-1}\,{}_2F_1(-n+2s-1, \tfrac{1}{2}(2s-1); 2s-1; 4pq) + a_{rs}n^{(s)}\binom{n-s}{s}p^{2s}q^{2s}\,{}_2F_1(-n+2s, \tfrac{1}{2}(2s+1); 2s+1; 4pq)\right\}. \quad (3\cdot21)$$

The following tables which will be useful in writing down explicitly the higher odd mean differences from either (3·19) or (3·21) show the numerical values of the coefficients a'_{rs} and a_{rs} for $r = 0(1)5$; $1 \leqslant s \leqslant r+1$.

Table 1. *Values of* a'_{rs}

r \ s	1	2	3	4	5	6
0	1	.	.	.	.	.
1	1	4	.	.	.	.
2	1	40	32	.	.	.
3	1	364	1,120	384	.	.
4	1	3,280	30,912	32,256	6,144	.
5	1	29,524	799,040	1,951,488	1,013,760	122,880

Table 2. *Values of* $a_{rs}/2^{2r}$

r \ s	1	2	3	4	5	6
0	1	.	.	.	.	.
1	1	1	.	.	.	.
2	1	5	2	.	.	.
3	1	21	28	6	.	.
4	1	85	294	180	24	.
5	1	341	2,816	3,762	1,320	120

4. The odd order mean differences for the Poisson distribution

The odd order mean difference Δ_{2r+1} for the Poisson distribution is given by

$$\Delta_{2r+1} = e^{-2\lambda}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}|i-j|^{2r+1}\frac{\lambda^i}{i!}\frac{\lambda^j}{j!} = 2e^{-2\lambda}\sum_{i=0}^{\infty}\sum_{j=0}^{i}(i-j)^{2r+1}\frac{\lambda^i}{i!}\frac{\lambda^j}{j!} \quad (4\cdot1)$$

for the same reason as for the binomial distribution.

The sums in (4·1) may be obtained in a straightforward manner; but this becomes quite tedious for $r > 2$. I have therefore considered the alternative simpler approach which is to obtain the limit of (3·21) for the binomial distribution as $n \to \infty$, $p \to 0$ and $np \to \lambda$.

Since $\lim_{\substack{n\to\infty,\ p\to 0\\ np\to\lambda}} {}_2F_1(-n+\alpha, \beta; \gamma; 4pq) = {}_1F_1(\beta; \gamma; -4\lambda)$, (3·21) under these limiting conditions becomes

$$\Delta_{2r+1} = 2\sum_{s=1}^{r+1}\left\{a'_{rs}\frac{\lambda^{2s-1}}{(2s-1)!}\,{}_1F_1(\tfrac{1}{2}(2s-1);\ 2s-1;\ -4\lambda) + a_{rs}\frac{\lambda^{2s}}{(2s)!}\,{}_1F_1(\tfrac{1}{2}(2s+1);\ 2s+1;\ -4\lambda)\right\}. \quad (4\cdot2)$$

Also, by virtue of the relation

$$\frac{\lambda^{\nu}}{\Gamma(\nu+1)}\,{}_1F_1(\tfrac{1}{2}(2\nu+1);\ 2\nu+1;\ -4\lambda) = e^{-2\lambda}\,I_{\nu}(2\lambda), \quad (4\cdot3)$$

where the function I_{ν} has been explained earlier, (4·2) becomes

$$\Delta_{2r+1} = 2e^{-2\lambda}\sum_{s=1}^{r+1}\lambda^s\{a'_{rs}I_{s-1}(2\lambda) + a_{rs}I_s(2\lambda)\}. \quad (4\cdot4)$$

It will be seen that the value of Δ_1 given by (1·4) agrees with the above, on letting $r = 0$; and

$$\left.\begin{aligned}\Delta_3 &= 2\lambda\, e^{-2\lambda}\{(1+4\lambda)\,I_0 + 4\lambda I_1\},\\ \Delta_5 &= 2\lambda\, e^{-2\lambda}\{(1+16\lambda+32\lambda^2)\,I_0 + 8\lambda(1+4\lambda)\,I_1\},\end{aligned}\right\} \quad (4\cdot5)$$

after successive reduction of $I_{\nu}(\nu \geqslant 2)$ to I_0 and I_1 by the recurrence relation

$$I_{\nu+1}(2\lambda) = I_{\nu-1}(2\lambda) - (\nu/\lambda)\,I_{\nu}(2\lambda). \quad (4\cdot6)$$

In (4·5), I_{ν} should be taken to mean $I_{\nu}(2\lambda)$. These values have also been checked independently by evaluating them directly from (4·1).

It is of some interest to consider the standardized ratio $\Delta_{2r+1}/(\text{variance})^{r+\frac{1}{2}}$, namely,

$$D_{2r+1} = \frac{2e^{-2\lambda}}{\lambda^{r+\frac{1}{2}}}\sum_{s=1}^{r+1}\lambda^s\{a'_{rs}I_{s-1}(2\lambda) + a_{rs}I_s(2\lambda)\}. \quad (4\cdot7)$$

A limiting form for (4·7) can be got for large λ by considering the asymptotic series for the modified Bessel function $I_{\nu}(2\lambda)$. This is given by

$$I_{\nu}(x) \simeq \frac{e^x}{\sqrt{2\pi x}}\left\{1 - \frac{4\nu^2-1^2}{1!\,8x} + \frac{(4\nu^2-1^2)(4\nu^2-3^2)}{2!(8x)^2} - \ldots\right\} + 0\left(\frac{e^{-x}}{\sqrt{x}}\right). \quad (4\cdot8)$$

This serves all practical purposes so long as ν is not of order $\sqrt{x}$ (Lehmer, 1943). Considering only the first term in (4·8), the limiting value of D_{2r+1} for large λ can be written as

$$D_{2r+1} \simeq \frac{2e^{-2\lambda}}{\lambda^{r+\frac{1}{2}}}\sum_{s=1}^{r+1}\lambda^s\frac{e^{2\lambda}}{\sqrt{2\pi}\sqrt{2\lambda}}(a'_{rs}+a_{rs}),$$

which on simplification and omission of terms of order λ^{-1}, λ^{-2}, etc. gives us

$$D_{2r+1} \simeq \frac{1}{\sqrt{\pi}}(a'_{r,r+1} + a_{r,r+1}). \quad (4\cdot9)$$

Substituting for the numerator from (3·18),

$$D_{2r+1} \simeq \frac{2}{\sqrt{\pi}}.2^{2r}.r!. \quad (4\cdot10)$$

When $r = 0$, the limiting value of D_1 equals $2/\sqrt{\pi}$, as demonstrated in one of the tables in my recent article (1958).

5. Conclusion

I may mention that although explicit results have been obtained here, it seems possible to derive certain interesting recurrence relations connecting the mean differences of different orders by considering the absolute moments of the distributions of the variate differences. This work is being pursued at the moment and I think I shall soon be able to record the outcome of this approach.

Finally, I wish to express my indebtedness to Prof. M. G. Kendall, under whose supervision the above investigations were carried out, for his continued support and encouragement.

REFERENCES

KAMAT, A. R. (1953). The third moment of Gini's mean difference. *Biometrika*, **40**, 451.

LEHMER, D. H. (1943). Note on the computation of the Bessel function $I_n(x)$. *Math. Tables and Aids to Computation*, **1**, 133.

LOMNICKI, Z. A. (1952). The standard error of Gini's mean difference. *Ann. Math. Statist.* **23**, 635.

NAIR, U. S. (1936). Standard error of Gini's mean difference. *Biometrika*, **28**, 428.

RAMASUBBAN, T. A. (1956). A χ-approximation to Gini's mean difference. *J. Indian Soc. Agric. Statist.* **8**, 116.

RAMASUBBAN, T. A. (1958). The mean differences and the mean deviations of some discontinuous distributions. *Biometrika*, **45**, 549.

Determination of parameters in the Johnson system of probability distributions

BY D. C. M. LESLIE

Royal Aircraft Establishment, Farnborough

The system of curves proposed by Johnson (1949) for the graduation of statistical distribution has been very successful in practice. The principal drawback is that it is not very easy to determine the four parameters defining the distribution from the observed moments. In this note, a method is suggested for calculating these parameters if the distribution under study is not far removed from normality. For a short summary of the Johnson system, the reader is referred to Draper (1952).

The Johnson system of curves is mainly used for fitting to empirical data, but it can also be used to graduate theoretical distributions. In the author's researches, the generating function of a distribution was obtained in closed form, but the inversion integral could not be evaluated. The first four moments could be found very easily, and this suggested the use of a system of frequency curves to fit the distribution approximately. It was found that the distribution was never very far from normality, and that it gave rise to β_1, β_2 points which always lay in the S_U region of the β_1, β_2 plane. Therefore, this note deals only with the fitting of curves of the S_U family to distributions which are not too non-normal.

The variable x has an S_U distribution if

$$z = \eta + \delta \sinh^{-1}\left(\frac{x-\xi}{\lambda}\right), \tag{1}$$

z being a unit normal variable.* The problem is to determine η, δ, ξ and λ given the first four moments μ_r of an observed distribution (this term including distributions which arise theoretically). From the opposite point of view, it is comparatively easy to calculate the μ_r given η, δ, ξ and λ. Formulae are given by Johnson (1949, p. 163), and from these formulae it follows that

$$\beta_1 = \omega(\omega-1)\frac{\{\omega(\omega+2)\sinh 3\Omega + 3\sinh\Omega\}^2}{2(\omega\cosh 2\Omega+1)^3} \tag{2}$$

and

$$\gamma_2 = \beta_2 - 3 = (\omega-1)\frac{\omega^2(\omega^3+3\omega^2+6\omega+6)\cosh 4\Omega + 4\omega(\omega+3)\cosh 2\Omega - 3(\omega-1)}{2(\omega\cosh 2\Omega+1)^2}, \tag{3}$$

where

$$\omega = e^{1/\delta^2} \quad \text{and} \quad \Omega = \eta/\delta. \tag{4}$$

The problem is now reduced to that of determining ω and Ω from given values of β_1 and γ_2; the functional form of equations (2) and (3) makes it clear that the problem is rather formidable.

However, it can be solved completely if the distribution is symmetrical about the origin, since then $\beta_1 = 0$, but $\gamma_2 \neq 0$. This implies $\Omega = 0$, so that equation (3) reduces to

$$\gamma_2 = \tfrac{1}{2}(\omega^2-1)(\omega^2+3)$$

and

$$\omega = \{\sqrt{(4+2\gamma_2)}-1\}^{\frac{1}{2}} = \{\sqrt{(2\beta_2-2)}-1\}^{\frac{1}{2}}. \tag{5}$$

* The letter η has been used in place of Johnson's γ because of the use of γ_2 for the standardized fourth cumulant.

This determines δ directly, while $\Omega = 0$ implies that $\eta = 0$ also. The remaining parameters ξ and λ are determined from the first two moments of x. Johnson's formulae are for the moments of the normalized variable

$$y = \frac{x-\xi}{\lambda}$$

and transforming them, we find

$$\mu_1' = \xi - \lambda\omega^{\frac{1}{2}}\sinh\Omega, \tag{6}$$

$$\mu_2 = \tfrac{1}{2}\lambda^2(\omega-1)(\omega\cosh 2\Omega + 1). \tag{7}$$

Since the distribution is symmetric, $\mu_1' = 0$, while it has been proved that $\Omega = 0$; thus $\xi = 0$ also, while

$$\lambda = \sqrt{\frac{2\mu_2}{\omega^2-1}}, \tag{8}$$

ω being given by (5).

When $\beta_1 \neq 0$, the problem does not seem to be soluble in terms of recognized functions. However, equations (2) and (3) show that if $\beta_1 = \gamma_2 = 0$ (i.e. if the distribution is normal) then $\omega = 1$. This suggests that if β_1 and γ_2 are small, an expansion of $\omega - 1$ in powers of β_1 and γ_2 may be useful; the existence of such an expansion follows from the analytical nature of the expressions (2) and (3). To simplify the work, we put

$$m = \omega\sinh^2\Omega \tag{9}$$

which reduces (2) and (3) to the purely algebraic forms

$$\beta_1 = (\omega-1)\frac{m[4(\omega+2)m + 3(\omega+1)^2]^2}{2(2m+\omega+1)^3}, \tag{10}$$

$$\gamma_2 = \frac{(\omega-1)P(\omega)}{2(2m+\omega+1)^2}, \tag{11}$$

where

$$P(\omega) = 8(\omega^3+3\omega^2+6\omega+6)m^2 + 8(\omega^4+3\omega^3+6\omega^2+7\omega+3)m + (\omega^5+3\omega^4+6\omega^3+10\omega^2+9\omega+3).$$

Expanding equations (10) and (11) about the point $\omega = 1$, we find

$$\beta_1 = \frac{9m}{m+1}(\omega-1) + O(\omega-1)^2, \tag{12}$$

$$\gamma_2 = \frac{4(4m+1)}{m+1}(\omega-1) + O(\omega-1)^2, \tag{13}$$

(12) and (13) may be inverted to give

$$\omega = 1 - \tfrac{1}{3}\beta_1 + \tfrac{1}{4}\gamma_2 + O(\beta_1^2, \text{etc.}), \tag{14}$$

$$m = \frac{4\beta_1}{9\gamma_2 - 16\beta_1} + O(\beta_1, \gamma_2). \tag{15}$$

Equation (14) in particular is attractive in its simplicity, and provided β_1 and γ_2 are small, this pair of equations enables η and δ to be determined very rapidly. It should be noted that β_1 and γ_2 must be known fairly accurately if good values are to be obtained for η and δ.

If β_1 and γ_2 are not small enough for (14) and (15) to give good results, adequate accuracy can sometimes be obtained by using more terms of these expansions. The process outlined above has, therefore, been extended to include terms of order $(\omega-1)^2$, and it is then found that

$$\omega = 1 - \tfrac{1}{3}\beta_1 + \tfrac{1}{4}\gamma_2 - \tfrac{31}{324}\beta_1^2 + \tfrac{11}{72}\beta_1\gamma_2 - \tfrac{1}{16}\gamma_2^2 + O(\beta_1^3), \tag{16}$$

$$m = \frac{4\beta_1}{9\gamma_2 - 16\beta_1} + \frac{(27\gamma_2^2 - 64\beta_1\gamma_2 + 40\beta_1^2)\beta_1}{6(9\gamma_2 - 16\beta_1)^2} + O(\beta_1^3). \tag{17}$$

If these expansions do not give adequate accuracy, they will at least provide useful starting values for a process of solving equations (10) and (11) by successive approximations; a considerable amount of labour would be needed to take them to the third order. Once η and δ have been found, ξ and λ may be determined from equations (6) and (7), which should be used in the forms

$$\mu_1' = \xi - \lambda m^{\frac{1}{2}}, \tag{18}$$

$$\mu_2 = \tfrac{1}{2}\lambda^2(2m+\omega+1). \tag{19}$$

It should be mentioned that Draper (1952) has solved the problem of inverting equations (2) and (3) heuristically. He has drawn extensive abacs of β_1 and γ_2 as functions of ω and Ω, and has then sought to fit empirical curves to these abacs. For ω he finds a power series not dissimilar to equation (16); he gives

$$\omega = 1 - 0{\cdot}312\beta_1 + 0{\cdot}227\gamma_2 - 0{\cdot}047\beta_1^2 + 0{\cdot}069\beta_1\gamma_2 - 0{\cdot}025\gamma_2^2 + O(\beta_1^3),$$

while (16) is equivalent to

$$\omega = 1 - 0{\cdot}333\beta_1 + 0{\cdot}250\gamma_2 - 0{\cdot}096\beta_1^2 + 0{\cdot}153\beta_1\gamma_2 - 0{\cdot}063\gamma_2^2 + O(\beta_1^3).$$

The coefficients of the first-order terms agree quite well, and those of the second-order terms have at least got the same signs. There is no very obvious correspondence between equation (17) and the various formulae given by Draper for Ω.

Dr Johnson has suggested a simple but instructive test of these formulae. A pair of values of ω and Ω is selected, and the corresponding values of β_1 and γ_2 calculated from equations (2) and (3); ω and Ω can then be estimated using the formulae given above, and the estimated values compared with the exact ones. Values obtained from equations (14) and (15) will be denoted by (ω_1, Ω_1) (first-order estimates) while those obtained from equations (16) and (17) will be denoted by (ω_2, Ω_2) (second-order estimates).

The table below shows the sort of results which are found when $(\omega - 1)$ is small

$\omega = 1{\cdot}1$	$\beta_1 = 0{\cdot}0103$	$\omega_1 = 1{\cdot}111$	$\omega_2 = 1{\cdot}0983$
$\Omega = 0{\cdot}1$	$\gamma_2 = 0{\cdot}4560$	$\Omega_1 = 0{\cdot}097$	$\Omega_2 = 0{\cdot}1001$
$\omega = 1{\cdot}1$	$\beta_1 = 0{\cdot}2182$	$\omega_1 = 1{\cdot}112$	$\omega_2 = 1{\cdot}0979$
$\Omega = 0{\cdot}5$	$\gamma_2 = 0{\cdot}7386$	$\Omega_1 = 0{\cdot}480$	$\Omega_2 = 0{\cdot}5011$
$\omega = 1{\cdot}1$	$\beta_1 = 0{\cdot}5758$	$\omega_1 = 1{\cdot}115$	$\omega_2 = 1{\cdot}0969$
$\Omega = 1{\cdot}0$	$\gamma_2 = 1{\cdot}2268$	$\Omega_1 = 0{\cdot}914$	$\Omega_2 = 0{\cdot}9985$
$\omega = 1{\cdot}3$	$\beta_1 = 0{\cdot}0400$	$\omega_1 = 1{\cdot}405$	$\omega_2 = 1{\cdot}2401$
$\Omega = 0{\cdot}1$	$\gamma_2 = 1{\cdot}6744$	$\Omega_1 = 0{\cdot}089$	$\Omega_2 = 0{\cdot}1037$

As long as γ_2 is not too large, the second-order formulae seem to give good results. Equation (5) shows that a series expansion in powers of γ_2 must diverge if γ_2 is greater than 2, and this implies that the method suggested above will only work well if $\gamma_2 < 1$. There must also be some limit on β_1. Equation (5) is, however, valid for all values of γ_2 as long as $\beta_1 = 0$. This suggests that an expansion of the form

$$\omega = f_0(\gamma_2) + \beta_1 f_1(\gamma_2) + \beta_1^2 f_2(\gamma_2) + \dots$$

might be very useful. It is already known that

$$f_0(\gamma_2) = \{\sqrt{(4 + 2\gamma_2)} - 1\}^{\frac{1}{2}}$$

and the next term should not be too difficult to calculate.

The author is indebted to Dr G. M. Jenkins for drawing his attention to the problem and for giving much helpful advice on how to solve it, and to Dr N. L. Johnson for suggesting the test of accuracy given above. Acknowledgment is made to the Controller, H.M. Stationery Office, for permission to publish this note.

REFERENCES

Draper, J. (1952). Properties of distributions resulting from certain simple transformations of the normal distribution. *Biometrika*, **39**, 290–301.

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika*, **36**, 149–76.

Note on a problem of estimation

By D. E. LLOYD

Ordnance Board, Ministry of Supply

Introduction

1. An estimation problem has arisen in which a random variable R from an unknown distribution with frequency function $p(r)$ has a probability of being observable which is a function of r denoted by $F(r)$. Thus the information available from an experiment consists of a set of observed values $r_1, r_2, \dots, r_m$,

together with the number $(n-m)$ of unobserved values. Truncated distributions are a special case which is frequently met in practice. However, in this note a case with a continuous function for $F(r)$ is considered.

2. In the examples considered, the information is used to estimate a parameter of the distribution function $p(r)$, $F(r)$ being assumed known. However, a similar method could be used to estimate parameters of $F(r)$. Estimates based on all the available information are compared with those based on $r_1, r_2, \ldots, r_m$ only, and with those based on the ratio m/n.

3. The chance of an observed value r is $p(r)\,F(r)$. Thus, if a single value of R is chosen at random, the chance that it will be observable is

$$P = \int p(r)\,F(r)\,dr \tag{1}$$

and the chance that it is not observable is $(1-P)$, the integral being taken over the whole range of r.

4. If a sample of size n is taken from the population, the probability of the observed values being $r_1, r_2, \ldots, r_m$ and there being $(n-m)$ unobserved values is

$${}^nC_m \prod_{i=1}^{m} \{p(r_i)\,F(r_i)\,dr_i\}\,(1-P)^{n-m}.$$

The log likelihood function is

$$L = \log {}^nC_m + \sum_{i=1}^{m} \log p(r_i) + \sum_{i=1}^{m} \log F(r_i) + (n-m)\log(1-P). \tag{2}$$

If θ is the parameter (or one of the parameters) of $p(r)$ to be estimated

$$\frac{\partial L}{\partial \theta} = \sum_{i=1}^{m} \frac{\partial}{\partial \theta}(\log p(r_i)) - \frac{n-m}{1-P}\frac{\partial P}{\partial \theta} = 0 \tag{3}$$

is the maximum likelihood equation for θ.

Estimate based solely on number of observed values

5. In a sample of size n, the probability of obtaining exactly m observed values is

$${}^nC_m P^m (1-P)^{n-m}.$$

The log likelihood function $L = m\log P + (n-m)\log(1-P)$ so that the maximum likelihood equation is

$$\frac{\partial L}{\partial \theta} = \frac{m}{P}\frac{\partial P}{\partial \theta} - \frac{n-m}{1-P}\frac{\partial P}{\partial \theta} = 0,$$

therefore

$$P(\hat{\theta}) = \frac{m}{n} \tag{4}$$

gives the required estimate.

Estimate based solely on $r_1, r_2, \ldots, r_m$

6. Given that m realizations of R are recorded, the probability that the values $r_1, r_2, \ldots, r_m$ will result is

$$\frac{\Pr\{r_1, r_2, \ldots, r_m \text{ and } (n-m) \text{ unrecorded values}\}}{\Pr\{m \text{ recorded and } (n-m) \text{ unrecorded values}\}},$$

$$= \frac{\prod_{i=1}^{m} \{p(r_i)\,F(r_i)\,dr_i\}\,(1-P)^{n-m}}{P^m(1-P)^{n-m}},$$

therefore

$$L = \sum_{i=1}^{m} \log p(r_i) + \sum_{i=1}^{m} \log F(r_i) - m\log P.$$

Thus

$$\frac{\partial L}{\partial \theta} = \sum_{i=1}^{m} \frac{\partial}{\partial \theta}\log p(r_i) - \frac{m}{P}\frac{\partial P}{\partial \theta} = 0 \tag{5}$$

gives the required estimate.

SPECIFIC FUNCTION FOR $p(r)$.

7. Now, put

$$p(r) = \frac{\alpha^{\frac{1}{2}\nu}}{\Gamma(\frac{1}{2}\nu)} r^{\frac{1}{2}\nu-1} e^{-\alpha r} \quad (0 \leqslant r < \infty).$$

This represents a χ^2 distribution with ν degrees of freedom and unknown dispersion.

Then, $\log p(r) = \frac{1}{2}\nu \log \alpha + (\frac{1}{2}\nu - 1)\log r - \alpha r - \log \Gamma(\frac{1}{2}\nu)$,

hence

$$\frac{\partial}{\partial \alpha} \log p(r) = \frac{\nu}{2\alpha} - r.$$

Substituting in equation (3), we find

$$\frac{m\nu}{2\alpha} - \sum_{i=1}^{m} r_i - \frac{n-m}{1-P} \frac{\partial P}{\partial \alpha} = 0. \tag{6}$$

Put

$$A = \frac{1}{m} \sum_{i=1}^{m} r_i, \quad B = \frac{n-m}{m}. \tag{7}$$

Then equation (6) becomes

$$A - \frac{\nu}{2\alpha} + \frac{B}{1-P} \frac{\partial P}{\partial \alpha} = 0. \tag{8}$$

Thus A and B are jointly sufficient statistics for the estimation of α.

Equations (4) and (5) give

$$P(\alpha) = \frac{m}{n} \tag{9}$$

and

$$A - \frac{\nu}{2\alpha} + \frac{1}{P} \frac{\partial P}{\partial \alpha} = 0 \tag{10}$$

for the other estimators of α.

SPECIFIC FUNCTION FOR $F(r)$

8. The function $F(r)$ must be chosen, as a result of prior information, to describe the conditions of the experiment. When truncation is assumed a function is chosen in the form

$$F(r) = 1 \quad (r \leqslant r_0),$$
$$= 0 \quad (r > r_0).$$

However, problems may arise in which, as a result of previous experiments, it seems more reasonable to use a continuous function for $F(r)$. We consider the case where

$$F(r) = \exp(-r/\lambda),$$

where λ is a constant, the value of which, it is assumed, is known.

9. For the particular functions chosen

$$P = \int_0^\infty F(r)\, p(r)\, dr$$
$$= \frac{\alpha^{\frac{1}{2}\nu}}{\Gamma(\frac{1}{2}\nu)} \int_0^\infty r^{\frac{1}{2}\nu-1} e^{-\{\alpha+(1/\lambda)\}r}\, dr$$
$$= \left(1 + \frac{1}{\alpha\lambda}\right)^{-\frac{1}{2}\nu}. \tag{11}$$

Hence

$$\frac{1}{1-P} \frac{\partial P}{\partial \alpha} = \frac{\nu}{2\alpha(\alpha\lambda+1)\left\{\left(1+\frac{1}{\alpha\lambda}\right)^{\frac{1}{2}\nu} - 1\right\}} \tag{12}$$

and

$$\frac{1}{P} \frac{\partial P}{\partial \alpha} = \frac{\nu}{2\alpha(\alpha\lambda+1)}. \tag{13}$$

Equations (12) and (13) can be used to substitute for P in equations (8) and (10), respectively.

THE CASE $\nu = 2$

10. This is the case where r is the squared radial distance of a point distributed in two dimensions in circular normal form.

Equations (11), (12) and (13) give

$$P = \left(1+\frac{1}{\alpha\lambda}\right)^{-1}, \tag{14}$$

$$\frac{1}{1-P}\frac{\partial P}{\partial \alpha} = \frac{1}{\alpha}\left(1+\frac{1}{\alpha\lambda}\right)^{-1} \tag{15}$$

and

$$\frac{1}{P}\frac{\partial P}{\partial \alpha} = \frac{1}{\alpha}(1+\alpha\lambda)^{-1}. \tag{16}$$

Using equation (15) to substitute for P in equation (8) gives

$$A\alpha^2 + \left(\frac{A}{\lambda} - 1 + B\right)\alpha - \frac{1}{\lambda} = 0. \tag{17}$$

The positive root of this equation gives the maximum likelihood estimate of α, denoted by $\hat{\alpha}_{ML}$.

11. Using equation (14) to substitute for P in equation (9) gives

$$1+\frac{1}{\alpha\lambda} = \frac{n}{m},$$

therefore

$$\alpha = \frac{1}{\lambda B}. \tag{18}$$

We denote this estimate of α by $\hat{\alpha}_B$

12. Using equation (16) to substitute into equation (10) gives

$$A - \frac{1}{\alpha} + \frac{1}{\alpha}(1+\alpha\lambda)^{-1} = 0.$$

Hence

$$\alpha = \frac{\lambda - A}{\lambda A}. \tag{19}$$

We denote this estimate of α by $\hat{\alpha}_A$

VARIANCES AND EFFICIENCIES OF ESTIMATES

13. For large n, the asymptotic variances of the three estimates are found to be

$$\operatorname{var}(\hat{\alpha}_{ML}) = \frac{\alpha^2}{n}(P - P^2 + P^3)^{-1}, \tag{20}$$

$$\operatorname{var}(\hat{\alpha}_B) = \frac{\alpha^2}{n}(P - P^2)^{-1}, \tag{21}$$

$$\operatorname{var}(\hat{\alpha}_A) = \frac{\alpha^2}{n}P^{-3}. \tag{22}$$

If we measure 'efficiency' in terms of inverse variance and regard α_{ML} as being 100% efficient, Fig. 1 shows how the large sample efficiencies of $\hat{\alpha}_A$ and $\hat{\alpha}_B$ vary with P. $\hat{\alpha}_A$ is more efficient when P is large, $\hat{\alpha}_B$ when P is small, as might be expected. The sum of the two efficiencies is always 100%.

OTHER VALUES OF ν

14. For a value of $\nu \neq 2$ the equation corresponding to equation (17) will not reduce to a quadratic equation in α. In these cases it is sometimes convenient to construct a chart plotting A/λ against B for fixed values of $\alpha\lambda$. From equations (8) and (12) it is evident that the graphs are all straight lines, and approximate values of $\hat{\alpha}_{ML}$ can easily be found using the chart.

15. In conclusion I wish to thank Mr D. J. Bishop for advice and encouragement.

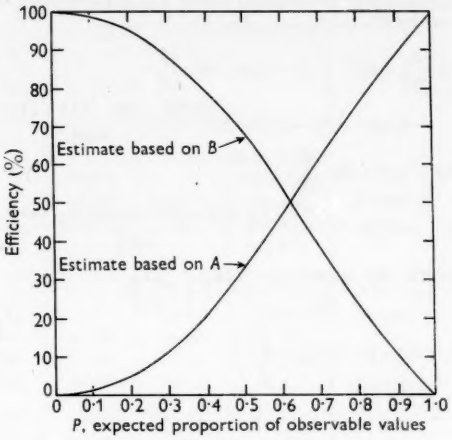


Fig. 1. Comparison of efficiencies.

On the non-central chi-square distribution

By MUNUSWAMY SANKARAN
Presidency College, Madras

1. Abdel-Aty (1954) has shown how the transformation

$$y = \left(\frac{\chi'^2}{f+\lambda}\right)^h = \left(\frac{\chi'^2}{r}\right)^h, \quad \text{where} \quad r = f+\lambda \tag{1}$$

may be used to obtain an approximation to the distribution of non-central χ^2, where f is the number of degrees of freedom, λ the non-central parameter and h an appropriate positive constant. To determine h, Aty examines the cumulants of y expressed in terms of cumulants of χ'^2, as power series in r^{-s}. He finds 'after expansion that $h = \frac{1}{3}$ is the appropriate value to choose for minimizing the coefficient of r^{-2} in $\kappa_3(y)$ and reducing the higher cumulants, $\kappa_s(y)$ for $s > 2$, to a minimum'. He then develops two methods of approximation.

(i) This assumes that $(\chi'^2/r)^{\frac{1}{3}}$ is normally distributed about $1-\frac{2}{9}(1+b)/r$ with variance $\frac{2}{9}(1+b)/r$, where $b = \lambda/(f+\lambda)$.

(ii) Alternatively, he suggests a 'closer approximation' which consists in applying the Cornish–Fisher (1937) expansion to the standardized deviate,

$$Y = \frac{y-\kappa_1(y)}{\sqrt{\{\kappa_2(y)\}}}, \tag{2}$$

where h is again taken as $\frac{1}{3}$.

2. In the present paper we shall use Aty's transformation (1) but instead of setting a fixed value to h, namely, $h = \frac{1}{3}$ here, we shall give h changing values so determined that the leading term in the third cumulant of y vanishes, thus following the procedure adopted by Wilson & Hilferty (1931). It follows that

$$h = 1-\frac{k_1 k_3}{3k_2^2} = 1-\frac{2}{3}\frac{(f+\lambda)(f+3\lambda)}{(f+2\lambda)^2}, \tag{3}$$

where

$$k_s = 2^{s-1}(s-1)!\,(f+s\lambda) \tag{4}$$

is the sth cumulant of χ'^2. It will be seen from (3) that $\frac{1}{3} \leqslant h \leqslant \frac{1}{2}$. We have for the first four cumulants of y, for the value of h given by (3),

$$\left.\begin{aligned}
\kappa_1(y) &= 1+h(h-1)\frac{k_2}{2k_1^2}+h(h-1)(h-2)(1-3h)\frac{k_2^2}{8k_1^4} \\
&\qquad +h(h-1)(h-2)(h-3)\frac{[2k_1^2k_4-(28-51h+11h^2)k_2^3]}{48k_1^6}+O\left(\frac{1}{r^4}\right), \\
\kappa_2(y) &= h^2\frac{k_2}{k_1^2}+h^2(h-1)(1-3h)\frac{k_2^2}{2k_1^4} \\
&\qquad +h^2(h-1)\frac{[(7h-11)k_1^2k_4-2(h-2)(35h^2-84h+40)k_2^3]}{12k_1^6}+O\left(\frac{1}{r^4}\right), \\
\kappa_3(y) &= h^3(1-h)\frac{[2(23-49h+23h^2)k_2^3-3k_1^2k_4]}{2k_1^6}+O\left(\frac{1}{r^4}\right), \\
\kappa_4(y) &= \frac{h^4}{3k_1^5}[3k_1k_4-4(4-5h)k_2k_3]+O\left(\frac{1}{r^4}\right), \\
\kappa_5(y) &= O\left(\frac{1}{r^4}\right)+\ldots.
\end{aligned}\right\} \tag{5}$$

3. Two approximations will now be considered.

(iii) *First approximation.* We take y as normally distributed with the following mean and standard deviation:

$$\left.\begin{aligned}
\mathscr{E}(y) &= 1+h(h-1)\frac{k_2}{2k_1^2}-h(h-1)(2-h)(1-3h)\frac{k_2^2}{8k_1^4}, \\
\sigma(y) &= \frac{hk_2^{\frac{1}{2}}}{k_1}\left\{1-\frac{(1-h)(1-3h)}{4k_1^2}k_2\right\}.
\end{aligned}\right\} \tag{6}$$

It should be noted that this approximation is similar to that given by Haldane (1937, p. 398) in his paper 'The approximate normalization of a class of frequency distributions'. We shall describe this as approximation H.

(iv) *Second approximation.** We shall in this case apply the Cornish–Fisher form of the Edgeworth expansion to the Y of equation (2), where h is as before given by (3) and the cumulants of y are obtained from (5), using the k_s of (4). It follows that considering terms up to order $r^{-\frac{3}{2}}$†, the 100α percentage point of the standardized deviate Y is given by

$$Y_\alpha = \xi_\alpha+\tfrac{1}{12}(\xi_\alpha^2-1)(1-h)\frac{[2(23-49h+23h^2)k_2^3-3k_1^2k_4]}{k_1^3k_2^3}+\tfrac{1}{72}(\xi_\alpha^3-3\xi_\alpha)\frac{[3k_1k_4-4(4-5h)k_2k_3]}{k_1k_2^2},$$

where ξ_α is the 100α percent point of a standardized normal deviate.

4. Some comparisons of approximations to the upper and lower 5 % points are made in Table 1 below for the same f and λ values used by Aty. It is seen that at the upper 5 % point the Edgeworth approximation (iv) of this paper, based on a varying h, is definitely more accurate than Aty's similar approximation (ii) with $h = \frac{1}{3}$. Further, the simpler approximation H, or (iii), is here somewhat better than Aty's Edgeworth approximation, and a great deal better than his simple normal approximation defined in (i) above (but not given in the table). At the lower 5 % point both Edgeworth approximations are good, but Aty's is somewhat the more accurate. The H approximation is not so good here, and is not shown in the table.

5. It is difficult to draw any general conclusions because exact values for other percentage points are not available for comparison, and it is by no means certain that an approximation which is best at one percentage point will also be so at another. The present investigation has, however, shown that under some conditions approximations based on a fixed value of $h = \frac{1}{3}$ will not be the best to use.

* I owe this suggestion to Dr F. N. David.

† Strictly this includes also the term $\frac{1}{2}\frac{1}{0}\frac{1}{0}k_5(y)\,k_2(y)^{-\frac{5}{2}}(\xi_\alpha^4-6\xi_\alpha^2+3)$. It is felt that omission of this as well as terms of $O(r^{-4})$ in k_1, k_2, k_3 and k_4 will not affect the conclusions reached here. The inclusion of these terms will only strengthen the conclusions, but will also considerably increase the computations involved.

Table 1. *Comparison of exact percentage points of the χ'^2 distribution with the values obtained from various approximations*

		Upper 5% points				Lower 5% points		
f	λ	Exact	Aty's 'closer approx.'	Approx. H	Edgeworth expansion to the present approx.	Exact	Aty's 'closer approx.'	Edgeworth expansion to the present approx.
2	1	8·642	8·38	8·58	8·656	0·168	0·171	0·157
	4	14·641	14·62	14·63	14·661	0·646	0·643	0·654
	16	33·054	33·08	33·07	33·060	6·322	6·320	6·331
	25	45·308	45·33	45·31	45·317	12·080	12·077	12·085
4	1	11·707	11·67	11·69	11·681	0·909	0·908	0·918
	4	17·309	17·27	17·28	17·318	1·765	1·766	1·763
	16	35·427	35·44	35·43	35·435	7·884	7·885	7·893
	25	47·613	47·62	47·62	47·621	13·733	13·732	13·737
7	1	16·004	15·99	15·98	16·013	2·494	2·494	2·497
	4	21·228	21·21	21·20	21·242	3·664	3·665	3·668
	16	38·970	38·96	38·96	38·978	10·257	10·256	10·264
	25	51·061	51·06	51·06	51·068	16·226	16·227	16·231

My thanks are due to Dr F. N. David for her helpful criticism of an earlier version of the paper. I also had the benefit of helpful suggestions of Prof. E. S. Pearson in the preparation of the paper for publication. Finally, I wish to record my appreciation to some of my students for computational work in connexion with the table.

REFERENCES

Abdel-Aty, S. H. (1954). *Biometrika*, **41**, 538–40.
Cornish, E. A. & Fisher, R. A. (1937). *Rev. Inst. Int. Statist.* **5**, 307–20.
Haldane, J. B. S. (1937). *Biometrika*, **29**, 392–404.
Wilson, E. B. & Hilferty, M. M. (1931). *Proc. Nat. Acad. Sci., Wash.*, **17**, 684.

On the distribution of the largest of six roots of a matrix in multivariate analysis

By K. C. Sreedharan Pillai* and Celia G. Bantegui†

The Statistical Center, University of the Philippines

1. Introduction

Roy (1945, 1953, 1957) has shown that tests of certain hypotheses and confidence interval estimation in multivariate situations can be based on those associated with the distribution of the extreme characteristic roots of a matrix whose elements are functions of the observations. The distribution of the extreme roots depends on the number of non-null roots, s, and on two other supplementary parameters, m and n, which are functions of the number of variates and, in some cases, the number of samples and the size of the samples.

* United Nations Senior Adviser in Mathematical Statistics and Visiting Professor of Statistics.
† Research Fellow, 1957–58, and presently Lecturer in Statistics.

Three different types of hypotheses are often posed for multivariate tests and the non-null characteristic roots of the matrix that arise in each test have a distribution, the form of which, under the null hypothesis, is the same for all the three tests. Roy (1939), Hsu (1939) and Fisher (1939) derived independently the distribution of the roots, $\theta_1, \ldots, \theta_s$, which can be given in the form

$$f(\theta_1, \ldots, \theta_s) = C(s, m, n) \prod_{i=1}^{s} \theta_i^m (1-\theta_i)^n \prod_{i>j} (\theta_i - \theta_j) \quad (\theta < \theta_1 \leqslant \ldots \leqslant \theta_s < 1), \tag{1}$$

where

$$C(s, m, n) = \frac{\pi^{\frac{1}{2}s} \prod_{i=1}^{s} \Gamma\left(\frac{2m+2n+s+i+2}{2}\right)}{\prod_{i=1}^{s} \Gamma\left(\frac{2m+i+1}{2}\right) \Gamma\left(\frac{2n+i+1}{2}\right) \Gamma\left(\frac{i}{2}\right)}, \tag{2}$$

and the parameters m and n are defined differently for the different types of hypotheses (Pillai, 1955).

In this paper, the cumulative distribution function of the largest of six roots has been studied with a view to tabulating the upper percentage points.

2. An approximation to the c.d.f. of the largest of six roots of a matrix

Pillai (1954) gave the explicit expression of the c.d.f. of θ_s for values of s ranging from two up to ten. Taking the case $s = 6$ and following his method (1954, 1956) an approximation to the c.d.f. of the largest of six roots, useful for computing the upper percentage points, has been obtained in the form:

$$\begin{aligned}
\Pr(\theta_6 \leqslant x) = 1 &- \frac{\Gamma(2m+2n+12)\,\Gamma(m+3)\,\Gamma(n+3)}{4\Gamma(m+n+4)\,\Gamma(2m+6)\,\Gamma(2n+6)} x^{m+5}(1-x)^{n+1} \\
&+ \frac{\Gamma(2m+2n+12)\,\Gamma(m+3)\,\Gamma(n+3)}{2(m+n+5)\,\Gamma(m+n+4)\,\Gamma(2m+5)\,\Gamma(2n+6)} x^{m+4}(1-x)^{n+1} \\
&- \frac{\Gamma(2m+2n+10)\,\Gamma(m+2)\,\Gamma(n+3)\,[3(m+n+5)(m+2)+(n+1)]}{2\Gamma(m+n+5)\,\Gamma(2m+4)\,\Gamma(2n+6)} x^{m+3}(1-x)^{n+1} \\
&+ \frac{\Gamma(2m+2n+10)\,\Gamma(m+4)\,\Gamma(n+3)}{2(m+2)\,\Gamma(m+n+5)\,\Gamma(2m+3)\,\Gamma(2n+6)} x^{m+2}(1-x)^{n+1} \\
&- \frac{\Gamma(2m+2n+8)\,\Gamma(m+4)\,\Gamma(n+3)}{2\Gamma(m+n+4)\,\Gamma(2m+3)\,\Gamma(2n+6)} x^{m+1}(1-x)^{n+1}.
\end{aligned} \tag{3}$$

For the details on the derivation of (3) the reader is referred to Bantegui (1958).

3. Upper percentage points

Pillai (1954, 1956, 1957) using his approximation formulae, gave the upper 5 and 1 % points for the largest root for $s = 2$ and $s = 5$. These values were given for $m = 0(1)4$ and n varying from 5 to 1000 and are given to three or four significant figures to ensure sufficient accuracy for interpolation. Similarly, Sen (1957) gave tables of the upper 5 and 1 % points for 3 roots and Ventura (1957) for 4 roots. Pillai has published all these four tables, for $s = 2$, 3, 4 and 5 (1957).

Foster & Rees (1957) have tabulated the upper percentage points (80, 85, 90, 95 and 99) of the largest root for $s = 2$, $m = -0{\cdot}5, 0(1)9$ and $n = 1(1)19(5)49, 59, 79$. Foster (1957, 1958) has further extended these tables for values of $s = 3$ and 4. The arguments they have used for tabulation are $\nu_1 = 2n+s+1$ and $\nu_2 = 2m+s+1$.

Bantegui (1958), also following the same approach as that of Pillai and using equation (3), obtained the percentage points for $s = 6$. These are presented in Tables 1 and 2. For computing the values of $\Pr(\theta_6 \leqslant x)$, a trial value was extrapolated from Pillai's tables (1957) of percentage points of the largest root for $s = 2$, 3, 4 and 5 for each of the $5 \times 16 = 80$ combinations of m and n; the trial value and a few values around it were then examined and backward linear interpolation carried out using two values, one giving probability above and the other below the significance level.

4. Error of approximation

The error of approximation of the upper percentage points of the distribution is taken to be the difference between the approximate and exact probabilities. Pillai (1954), Sen (1957) and Ventura (1957) made such

Table 1. *Upper 5% points of the largest root for $s = 6$*

n	m				
	0	1	2	3	4
5	0·8246	0·8499	0·8685	0·8830	0·8945
10	·6552	·6917	·7206	·7442	·7639
15	·5371	·5758	·6077	·6346	·6577
20	·4535	·4912	·5231	·5505	·5746
25	·3918	·4276	·4583	·4852	·5091
30	0·3447	0·3782	0·4074	0·4332	0·4564
40	·2775	·3069	·3329	·3563	·3776
60	·1995	·2225	·2433	·2624	·2801
80	·1556	·1745	·1916	·2075	·2224
100	·1275	·1434	·1580	·1716	·1843
130	0·10036	0·11319	0·12504	0·13615	0·14666
160	·08272	·09348	·10388	·11284	·12175
200	·06702	·07586	·08409	·09186	·09926
300	·04545	·05156	·05728	·06281	·06790
500	·02765	·03143	·03498	·03835	·04160
1000	·01397	·01590	·01772	·01946	·02113

Table 2. *Upper 1% points of the largest root for $s = 6$*

n	m				
	0	1	2	3	4
5	0·8745	0·8929	0·9065	0·9169	0·9255
10	·7173	·7482	·7724	·7922	·8086
15	·5986	·6334	·6619	·6858	·7063
20	·5111	·5462	·5757	·6010	·6231
25	·4450	·4790	·5081	·5335	·5559
30	0·3936	0·4261	0·4542	0·4789	0·5011
40	·3194	·3484	·3739	·3969	·4177
60	·2315	·2548	·2757	·2948	·3125
80	·1814	·2006	·2181	·2342	·2493
100	·1491	·1654	·1803	·1942	·2072
130	0·11762	0·13091	0·14314	0·15457	0·16536
160	·09713	·10830	·11901	·12834	·13754
200	·07880	·08803	·09659	·10466	·11232
300	·05355	·05996	·06594	·07160	·07701
500	·03270	·03661	·04034	·04388	·04727
1000	·01651	·01855	·02046	·02229	·02405

studies for $s = 2$, 3 and 4, respectively. A comparative study of the values of the error of approximation they have obtained indicates that;

(1) There is greater agreement between the probabilities for the approximate and exact cases in the upper 1 % points than in the 5 %.

(2) The difference between the approximate and the exact probabilities in the upper 5 % points occur in the fifth decimal place, which on rounding gives a difference of only one in the fourth decimal place.

(3) Taking any value of the parameter m constant, the error of approximation increases slowly as the other parameter is increased; such increase occurs only in the sixth decimal place or at most is unity in the fifth decimal place when rounding.

The error of approximation was similarly studied for $s = 6$. From Table 1, the percentage point for the parameter $m = 0$, $n = 500$ is 0·02765. At this percentage point, the numerical value of the approximate C.D.F. is 0·9500515, while the probability based on the exact formula is calculated to be 0·9500105. The error of approximation is then equal to 0·0000410 which occurs in the fifth decimal place. This result is consistent with that for lower values of s. It can therefore be concluded that the approximation formula (3) is sufficiently accurate for the computation of the upper 5% or lower percentage points of the C.D.F. of the largest of six roots.

It is to be noted that the error of approximation discussed above is not the same as the error in the percentage points. However, a comparison of the percentage points obtained, for instance, for $s = 3$ based on Pillai's approximate formula (Pillai, 1957) with Foster's table (1957) reveals that the approximation is sufficiently accurate to justify the retention of 4 and 5 decimal places for the 95 % level. In fact, the approximate formula is adequate even for percentage points slightly below the 95% level. Foster & Rees (1957) have also discussed this point for $s = 2$. Since the error of approximation obtained for $s = 6$ agrees with that computed for smaller values of s, the accuracy of the approximation for the present case is evident.

5. Summary

In this paper the cumulative distribution function of the largest of six roots of a matrix in multivariate analysis has been studied. An approximation for obtaining upper percentage points (5 % or less) to the C.D.F. of the largest of six roots has been obtained following Pillai's method and numerical comparison was made of the probabilities obtained using the exact C.D.F. and the approximation to it. Since the comparison showed that the error of approximation was negligible, upper 5 and 1 % points were computed for small integral values of one parameter and values of a second parameter ranging up to 1000.

The authors wish to thank the Statistical Center, University of the Philippines for the facilities extended them during the preparation of this study.

REFERENCES

Bantegui, C. G. (1958). 'On the distribution of the largest of six roots of a matrix in multivariate analysis.' Unpublished Thesis, The Statistical Center, University of the Philippines.

Fisher, R. A. (1939). *Ann. Eugen., Lond.*, **9**, 238.

Foster, F. G. (1957). *Biometrika*, **44**, 441.

Foster, F. G. (1958). *Biometrika*, **45**, 492.

Foster, F. G. & Rees, D. H. (1957). *Biometrika*, **44**, 237.

Hsu, P. L. (1939). *Ann. Eugen., Lond.*, **9**, 250.

Pillai, K. C. S. (1954). 'On some distribution problems in multivariate analysis.' *Mimeo. Series* No. 88, Institute of Statistics, University of North Carolina.

Pillai, K. C. S. (1955). *Ann. Math. Statist.* **26**, 117.

Pillai, K. C. S. (1956). *Biometrika*, **43**, 122.

Pillai, K. C. S. (1957). *Concise Tables for Statisticians.* The Statistical Center, University of the Philippines.

Roy, S. N. (1939). *Sankhyā*, **4**, 381.

Roy, S. N. (1945). *Sankhyā*, **7**, 133.

Roy, S. N. (1953). *Ann. Math. Statist.* **24**, 220.

Roy, S. N. (1957). *Some Aspects of Multivariate Analysis.* New York: John Wiley and Sons, N.Y.

Sen, P. (1957). 'On a multivariate test criterion and its applications.' Unpublished Thesis, The Statistical Center, University of the Phillippines.

Ventura, S. R. (1957). 'On the extreme roots of a matrix in multivariate analysis and associated tests.' Unpublished Thesis, The Statistical Center, University of the Philippines.

The rank analogue of product-moment partial correlation and regression, with application to manifold, ordered contingency tables†

BY ROBERT H. SOMERS

Bureau of Applied Social Research, Columbia University

Daniels (1944) has demonstrated that the ordinary product-moment correlation, Spearman's rho, and Kendall's tau, are all special cases of a generalized correlation coefficient. This paper will demonstrate that a generalized *partial* correlation coefficient can be defined, as a simple extension of Daniels's theory. Kendall's partial tau will be seen to be a special case of this product-moment type partial coefficient. Thus, it is neither coincidental, nor remarkable (Kendall, 1955; Moran, 1950), that Kendall's partial tau is the same function of total taus that the product-moment partial is of total r. While the partial coefficient described here is of the first order, its form suggests a generalization to partial correlation of any order, which will be investigated at a later date. The theory will be seen to be immediately applicable to partial association in manifold, ordered, contingency tables, inasmuch as the partial tau here described is equally applicable to the case of ties among the rankings. As a further extension of product-moment theory to order statistics, the regression equations implied by Kendall's τ_a and τ_b will be described and illustrated, and the possibility of partial regression indicated.

No remarks are made here about sampling distributions. However, it is hoped that the generalized partial coefficient defined below will contribute to the difficult problem of the sample distribution of partial tau.

Daniels's generalized coefficient of correlation,

$$\Gamma = \frac{\Sigma a_{ij} b_{ij}}{\sqrt{\{\Sigma a_{ij}^2 \Sigma b_{ij}^2\}}},$$

is constructed in the following manner (cf. also Kendall, 1955, Ch. 2). When considering the correlation over n objects between two variables, P and Q, the objects are first numbered in arbitrary order, and all n^2 permuted pairs of objects are formed. Each pair is assigned a certain score, a_{ij}, for its value on variable P, and b_{ij} for variable Q. The ordinary product moment r becomes a special case of Γ when the assigned score, a_{ij}, is the difference in variate values of the ith and jth objects. Spearman's rho becomes a special case when the difference in ranking, rather than difference in variate value, is the score assigned to each pair. Finally, Kendall's tau is a special case of Γ when each pair is assigned a score of $\{+1/-1\}$ for pairs in which the jth member of the pair ranks {higher/lower} than the ith member, and 0 when the ith object is paired with itself, or when the ith and jth objects have the same rank.

This last case gives Kendall's τ_b. Since manifold, ordered, contingency tables may be considered rankings of objects with many ties, Kendall's τ_b is a useful coefficient of correlation for such data.

For convenience, we shall change Daniels's notation. We shall designate the scores as u_i for variable P, v_i for variable Q, and w_i for variable R, for $i = 1, 2, \ldots, q$, where $q = n^2$. The manner of assigning scores is unchanged; we shall simply let Daniels' a_{ij} become $u_{n(i-1)+j}$, for example. Thus Daniels's generalized coefficient becomes

$$\Gamma_{PQ} = \frac{\Sigma u_i v_i}{\sqrt{\{\Sigma u_i^2 \Sigma v_i^2\}}} \quad (\text{for } i = 1, \ldots, q).$$

In the following discussion we shall consider the u_i, v_i and w_i scores as components of vectors in q-dimensional Euclidian space, always taken relative to a fixed, normal, orthogonal basis. Thus, if we have n objects ranked on variables P, Q and R, we shall use the symbols $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$ to represent vectors having $q = n^2$ components,

$$\mathbf{P} = (u_1, u_2, \ldots, u_q), \quad \mathbf{Q} = (v_1, v_2, \ldots, v_q), \quad \mathbf{R} = (w_1, w_2, \ldots, w_q).$$

Further, as in ordinary vector analysis

$$|\mathbf{P}| = \sqrt{(u_1^2 + u_2^2 + \ldots + u_q^2)},$$

† This article may be identified as Publication No. A–252 of the Bureau of Applied Social Research, Columbia University. I am indebted to my colleague, Herbert Menzel, whose analytical problem led to, and whose ingenuity aided in, the pursuit of these results. The preparation of this paper was facilitated by funds obtained from the Bureau of Applied Social Research.

and we shall designate the unit vectors

$$\alpha_1 = \frac{\mathbf{P}}{|\mathbf{P}|}, \quad \alpha_2 = \frac{\mathbf{Q}}{|\mathbf{Q}|}, \quad \alpha_3 = \frac{\mathbf{R}}{|\mathbf{R}|}.$$

The scalar ('inner' or 'dot') product of two vectors is, as usual, the scalar

$$(\mathbf{P}.\mathbf{Q}) = u_1v_1 + u_2v_2 + \ldots + u_qv_q.$$

At this point we diverge from the usual vector notation. We need to employ the notion of vector (or 'outer') product. But this is ordinarily defined only in three-dimensional space, and we operate in a space of far higher dimension. The generalization of the vector product to such a space ordinarily requires tensor notation (Kaplan, 1953, p. 161), but we shall avoid introducing new notation by defining a generalized vector product, which can be treated as a vector in a companion space, just as it is in 3-space;

$$(\mathbf{P} \times \mathbf{Q}) = (u_1v_2 - u_2v_1, u_1v_3 - u_3v_1, \ldots, u_{q-1}v_q - u_qv_{q-1}),$$

i.e. a vector whose components are determinants of the 2×2 matrices of the form

$$\begin{vmatrix} u_i & u_j \\ v_i & v_j \end{vmatrix}, \quad \text{all } i < j, \text{ lexicographic ordering.}$$

This vector therefore has $\frac{1}{2}q(q-1)$ components, and is the vector perpendicular to the plane defined by the vectors $\mathbf{P}$ and $\mathbf{Q}$. (Note that $(\mathbf{P} \times \mathbf{Q}) = -(\mathbf{Q} \times \mathbf{P})$.)

Having established these definitions, we see immediately that

$$\tau_{PQ} = (\alpha_1.\alpha_2), \quad \tau_{PR} = (\alpha_1.\alpha_3) \quad \text{and} \quad \tau_{QR} = (\alpha_2.\alpha_3).$$

We shall show that

$$\tau_{QR.P} = \frac{(\alpha_1 \times \alpha_2).(\alpha_1 \times \alpha_3)}{|(\alpha_1 \times \alpha_2)|\,|(\alpha_1 \times \alpha_3)|} \tag{1}$$

is precisely the partial tau defined by Kendall as a phi coefficient† of a certain fourfold table (Kendall, 1955, p. 118), when the vectors α_i consist of scores appropriate to tau. With the appropriate scoring, this apparently becomes either a partial Spearman's rho, or the product-moment partial r.

From the analogous theory in 3-dimensional vector space, this form suggests that the partial can be treated as the cosine of the angle between the perpendiculars to the planes defined by the vectors α_1, α_2 and α_1, α_3. Geometrically, this is seen to be equivalent to the representation of the ordinary product moment partial coefficient (cf. Kendall, 1948, p. 372) as the cosine of a different, but equal, angle—the angle between the two planes. In the ordinary representation of orthogonal projections on a subspace of one less dimension, the partial would be expressed as (cf. Birkhoff & MacLane (1949), p. 187)

$$\tau_{QR.P} = \frac{[\alpha_2 - (\alpha_1.\alpha_2)\alpha_1].[\alpha_3 - (\alpha_1.\alpha_3)\alpha_1]}{|\alpha_2 - (\alpha_1.\alpha_2)\alpha_1|\,|\alpha_3 - (\alpha_1.\alpha_3)\alpha_1|},$$

which is equivalent to the above form. However, the former has obvious advantages of representation, and immediately suggests the generalization to a partial of any order, since the vector product can be generalized to any number of vectors.

In order to show that (1) is the partial tau defined by Kendall, it is only necessary to observe that

$$(\alpha_1 \times \alpha_2).(\alpha_1 \times \alpha_3) = (\alpha_1.\alpha_1)(\alpha_2.\alpha_3) - (\alpha_1.\alpha_3)(\alpha_1.\alpha_2).$$

This is a well-known identity in 3-space, and will be demonstrated in the general case below. If this is true, then, since the α_i are unit vectors, $(\alpha_1.\alpha_1) = 1$. Moreover, $(\alpha_2.\alpha_3) = \tau_{QR}$, $(\alpha_1.\alpha_2) = \tau_{PQ}$ and $(\alpha_1.\alpha_3) = \tau_{PR}$. Thus the numerator is, as it should be, $\tau_{QR} - \tau_{PQ}\tau_{PR}$. In the denominator, we have

$$|(\alpha_1 \times \alpha_2)| = \sqrt{[(\alpha_1 \times \alpha_2).(\alpha_1 \times \alpha_2)]} = \sqrt{[(\alpha_1.\alpha_1)(\alpha_2.\alpha_2) - (\alpha_1.\alpha_2)(\alpha_1.\alpha_2)]} = \sqrt{(1 - \tau_{PQ}^2)};$$

and similarly for $|(\alpha_1 \times \alpha_3)|$. We have thus reduced the vector definition to Kendall's partial tau. Furthermore, the form of the partial immediately suggests the partial regression coefficients and partial covariance. It might be noted that the partial variance $(|\alpha_1 \times \alpha_2|)^2$ is the determinant $|A_2A_2'|$, where A_2 is the $2 \times q$ matrix with rows α_1 and α_2.

† The phi coefficient, $\dfrac{ad - bc}{\sqrt{\{(a+b)(a+c)(b+d)(c+d)\}}}$, it should be noted, is the product-moment correlation of a fourfold table.

We shall show that this equivalence between partial τ as the cosine of the angle between vectors, and partial τ as defined by Kendall, holds by enumerating the vector product. We have:

$$\alpha_1 = (u_1', u_2', \ldots, u_q'),$$
$$\alpha_2 = (v_1', v_2', \ldots, v_q'),$$
$$\alpha_3 = (w_1', w_2', \ldots, w_q'),$$

i.e. unit vectors, where, for example, $u_i' = u_i/|\mathbf{P}|$. Kendall's partial τ is

$$\tau_{QR.P} = \frac{\Sigma v_i' w_i' - \Sigma u_i' v_i' \Sigma u_i' w_i'}{\sqrt{\{[1-(\Sigma u_i' v_i')^2][1-(\Sigma u_i' w_i')^2]\}}}.$$

We will show that the numerator of this coefficient is equivalent to $(\alpha_1 \times \alpha_2).(\alpha_1 \times \alpha_3)$. Since we used the same expansion rule for vector products in establishing the denominator, this will suffice for establishing the equivalence.

By definition (and omitting primes on u_i', v_i' and w_i'),

$$(\alpha_1 \times \alpha_2).(\alpha_1 \times \alpha_3) = (u_1v_2 - u_2v_1)(u_1w_2 - u_2w_1) + \ldots + (u_1v_q - u_qv_1)(u_1w_q - u_qw_1)$$
$$+ (u_2v_3 - u_3v_2)(u_2w_3 - u_3w_2) + \ldots + (u_{q-1}v_q - u_qv_{q-1})(u_{q-1}w_q - u_qw_{q-1}).$$

This can be broken up into some positive terms with $u_i^2 v_j w_j$, and some negative terms of the form $u_iu_j(v_iw_j + v_jw_i)$, the former for $i \neq j$, the latter for $i < j$. These two groups of terms are reduced as follows:

$$\Sigma u_i^2 v_j w_j = \sum_i u_i^2 \sum_j v_j w_j - \Sigma u_i^2 v_i w_i \quad (i \neq j, \quad \text{all} \quad i,j = 1, \ldots, q)$$

$$\sum_{i<j} u_iu_j(v_iw_j + v_jw_i) = \tfrac{1}{2}[\sum_i u_i \sum_j u_j(v_iw_j + v_jw_i) - 2\Sigma u_i^2 v_i w_i] \quad (\text{all } i,j).$$

But, since these are unit vectors, $\Sigma u_i^2 = 1$. Thus

$$(\alpha_1 \times \alpha_2).(\alpha_1 \times \alpha_3) = \Sigma v_i w_i - \tfrac{1}{2}[\Sigma u_iu_j(v_iw_j + v_jw_i)] \quad (\text{all } i,j),$$

and we merely need to show that the negative term is $\Sigma u_iv_i \Sigma u_iw_i$. Since the matrix of terms $u_iu_j(v_iw_j + v_jw_i)$ is symmetric

$$\tfrac{1}{2}[\Sigma u_iu_j(v_iw_j + v_jw_i)] = \sum_{i<j} u_iu_j(v_iw_j + v_jw_i) + \Sigma u_i^2 v_i w_i \quad (\text{all } i,j)$$
$$= \sum_{i<j} u_i \Sigma u_j(v_iw_j) + \sum_{i<j} u_i \Sigma u_j(v_jw_i) + \Sigma u_i^2 v_i w_i = \Sigma u_iv_i \Sigma u_iw_i,$$

and the equivalence is established.

A numerical example of the partial τ computed by means of the vector formulae will serve for illustration. For simplicity we shall avoid the redundancy involved in τ when computed from Daniels's generalized coefficient. As Kendall (1955, p. 20) notes, the Daniels's scoring rules give a numerator and denominator of τ twice the size of those given by the usual computing methods. Similarly, Daniels's notation increases the size of the vectors we have to consider, by adding redundant information. Since these vectors are large in any case, we have computed by means of reduced vectors, those consisting only of terms a_{ij} such that $i < j$.

Using the example of Kendall (1955, p. 117), and denoting the reduced vectors by the superscript,* we see that **P**, **Q** and **R** each have 15 components; the scores assigned to the 15 pairs as shown at the bottom of p. 117. Obviously $|\mathbf{P}^*| = |\mathbf{Q}^*| = |\mathbf{R}^*| = \sqrt{15}$, and the α_i are unit vectors constructed in the usual fashion. Then $(\alpha_1 \times \alpha_2)^*$, $(\alpha_1 \times \alpha_3)^*$ and $(\alpha_2 \times \alpha_3)^*$ are vectors with $\frac{1}{2}m(m-1)$, or 105, components, where $m = \frac{1}{2}n(n-1)$, n being the number of objects ranked. Each component is a score assigned to each distinguishable *pair* of *pairs*, the score being $\pm 2/m$ or 0. In the tied case there will also be scores of $\pm 1/m$. In this example, we find

$$|(\alpha_1 \times \alpha_2)^*| = \sqrt{\{15^{-2}[4(25+19)]\}} = \sqrt{\frac{176}{15^2}},$$

$$|(\alpha_1 \times \alpha_3)^*| = \sqrt{\{15^{-2}[4(35+19)]\}} = \sqrt{\frac{216}{15^2}},$$

and

$$(\alpha_1 \times \alpha_2)^*.(\alpha_1 \times \alpha_3)^* = \frac{4(6-15)}{15^2},$$

giving $$\tau_{QR.P} = \frac{-36}{\sqrt{\{(216)(176)\}}} = -0{\cdot}185,$$

in agreement with Kendall.

For those who work with manifold, ordered, contingency tables, it should be emphasized that these vector operations hold equally well for τ_b. Thus, when there are tied ranks

$$\tau_{b_{QR.P}} = \frac{\tau_{b_{QR}} - \tau_{b_{PQ}}\tau_{b_{PR}}}{\sqrt{\{(1-\tau^2_{b_{PQ}})\,(1-\tau^2_{b_{PR}})\}}}$$

is also the product-moment analogue of a partial correlation for contingency tables.†

As illustration of the meaning of Kendall's tau, it may be useful to consider the regression equations implied by the product-moment analogy. From our definition of $\tau_{QR} = (\alpha_2 . \alpha_3) = \dfrac{(\mathbf{Q}.\mathbf{R})}{|\mathbf{Q}|\,|\mathbf{R}|}$, it is clear that the variance of scores on Q may be defined as $|\mathbf{Q}|^2$. Thus, regression coefficients may be defined in the same way as in product-moment theory. In the case of total regression, for variables P and Q with no ties

$$\beta_1 = \beta_2 = \frac{(\mathbf{P}.\mathbf{Q})}{|\mathbf{Q}|^2} = \frac{(\mathbf{P}.\mathbf{Q})}{|\mathbf{P}|^2},$$

since $|\mathbf{P}| = |\mathbf{Q}| = n(n-1)$, where we have ranked n objects (without ties). The same example given by Kendall may be used to illustrate these coefficients. Here we have six objects ranked without ties, on variables P and Q. There are thus 36 possible pairs (considering all permutations) and six of them will be assigned a score of 0 on both variables, since they are of the form a_{ii}. According to the scoring described above, and following Daniels, these pairs give the following scatterplot.

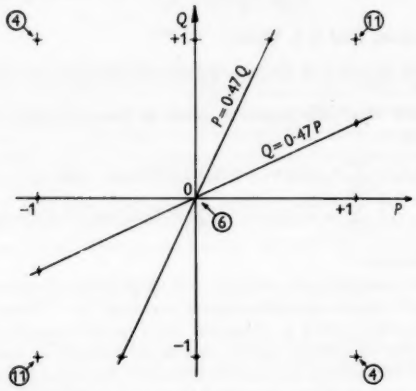


Fig. 1. Scatter diagram of pairs of observations implied by Kendall's τ_a.

(The circled figures are the number of permuted pairs given the indicated P and Q scores.)

$$\beta_1 = \beta_2 = \frac{(\mathbf{P}.\mathbf{Q})}{n(n-1)} = \frac{14}{30},$$

$\overline{p} = \overline{q} = 0$, so the regression equations are as shown,

$P = 0{\cdot}47Q$ and $Q = 0{\cdot}47P$,

$\tau_{PQ} = \sqrt{(\beta_1\beta_2)} = 7/15$, in agreement with Kendall.

In the case of tied rankings, say a contingency table as shown in Table 1 (from Lipset *et al.* (1951), p. 347) the scatterplot of pairs would be that given in Fig. 2.

† Such a partial coefficient for the tied case was used in an illustrative example, without justification, in a recent text; Sidney Siegel, *Nonparametric Statistics for the Behavioral Scientist*, p. 227. New York, 1956.

Table 1

		Level of knowledge			
		Low	Medium	High	Totals
Voting in accordance with predisposition	Yes	46	55	39	140
	No	9	35	38	82
	Totals	55	90	77	222

Since this is equivalent to ranking 222 objects, with many ties, there will be 49,284 permuted pairs entering into the scatterplot, in the following manner, where Q represents knowledge and R, vote.

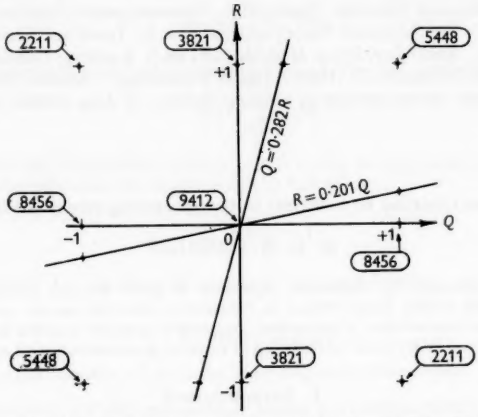


Fig. 2. Scatter diagram of pairs of observations implied by Kendall's τ_b.

(Again, the circled numbers are the number of permuted pairs at the given points.)

From the scatterplot, $\beta_1 = \dfrac{3237}{16115} = 0{\cdot}201, \quad \beta_2 = \dfrac{3237}{11480} = 0{\cdot}282$

Again the means are 0, and the regression equations become,

$$Q = 0{\cdot}201\,R, \quad R = 0{\cdot}282\,Q.$$

The value of $\sqrt{(\beta_1\beta_2)} = \dfrac{3237}{\sqrt{\{(16115)\,(11480)\}}} = 0{\cdot}238$. This, as we would expect, is exactly the value of τ_b for the contingency table given above. It will be noted that only in the case of ties among the rankings is it possible for the regression coefficients to be unequal. This is obvious from the redundancies in the scatterplot. It will be noticed that the first row of the scatterplot (Fig. 2) is the reverse of the last, and similarly for the columns. Again, this redundancy enters in because of the rule used in assigning scores; $a_{ij} = -a_{ji}$.

Further work is in progress on the order statistics analogues of product-moment statistics. On the basis of Daniels's initial conceptualization, it seems clear that a general partial and multiple correlation and regression theory for order statistics can be generated. Similarly, the correlation ratios, for the case of tied

ranks, seem to be equivalent to assymetric coefficients of association for a contingency table that does not have ordered variables, i.e. attaining their maxima when the observations in any row or column fall entirely within one cell.

Finally, we may suggest an alternative point of view to those who reject such product moment derivatives as phi on the grounds that they have no 'operational meaning' (Goodman & Kruskal, 1954). While such meaning for a coefficient is often desirable, there are other times when *theoretical power* assumes greater importance. Often both theoretical power and operational meaning cannot be attained. The present investigations have resulted from a need for a coefficient of association in manifold, ordered, contingency tables which would permit logical inferences to be made about operationally undefined intervening variables. Thus a coefficient was needed which would give information about a partial coefficient when only the total coefficients were known. Obviously, if a choice had to be made in this case, theoretical power could not be sacrificed for the sake of operational meaningfulness.

REFERENCES

Birkhoff, G. & MacLane, S. (1949). *A Survey of Modern Algebra.* New York: The Macmillan Company.

Daniels, H. E. (1944). The relation between measures of correlation in the universe of sample permutations. *Biometrika*, **33**, 129–35.

Goodman, L. & Kruskal, Wm. (1954). Measures of association for cross classification. *J. Amer. Statist. Assoc.* **49**, 732–64.

Kaplan, W. (1953). *Advanced Calculus.* Cambridge. Massachusetts: Addison Wesley.

Kendall, M. G. (1948). *The Advanced Theory of Statistics*, **1**. London: Charles Griffin and Co. Ltd.

Kendall, M. G. (1955). *Rank Correlation Methods* (2nd ed.). London: Charles Griffin and Co. Ltd.

Lipset, S. M., Trow, M. Coleman, J. (1956). *Union Democracy.* Glencoe Illinois: Free Press.

Moran, P. (1950). Recent developments in ranking theory. *J. Roy. Statist. Soc.* Ser. B, **12**, 153–62.

The busy period in relation to the queueing process GI/M/1

By B. W. CONOLLY

ABSTRACT. A treatment by difference equations is given of the probability distributions associated with length of the Busy Period in connexion with the simple queueing process with general input and random service. A numerical comparison is made in a few typical cases between the processes $M/M/1$ and $D/M/1$, the latter being of interest in connexion with appointment systems.

1. Introduction

Kendall (1953) suggested the notation $GI/G/N$ as a compact description of an N-server queueing process in which the inter-arrival intervals are independently and arbitrarily distributed, and the service intervals are also arbitrarily distributed. 'GI' refers to input, and 'G' to service.

The process which will be considered here has the general independent input, while the service has probability density function (p.d.f.) of the negative exponential, or 'Markovian', type, viz.

$$e^{-t/b}/b. \tag{1·1}$$

Service is supplied by one server, and the queue is dealt with in order of arrival. Inter-arrival intervals and service intervals are statistically independent.

A consequence of the negative exponential assumption for service is that the probability $b_n(t)$ of the completion of n consecutive service periods in an interval $(0, t)$, the $(n+1)$th being in progress at the instant t, is given by

$$b_n(t) = e^{-t/b}(t/b)^n/n!. \tag{1·2}$$

Denoting the p.d.f. of inter-arrival intervals by

$$a(t) \quad (0 \leqslant t < \infty) \tag{1·3}$$

with mean a, we shall indicate by $c_n(t)$ the p.d.f. associated with an inter-arrival interval of length t during which time n consecutive service periods are completed, the $(n+1)$th being in progress at the end of the interval. Thus,

$$c_n(t) = a(t)\, b_n(t). \tag{1·4}$$

Most of the following investigation will involve Laplace transforms. In general, we shall use the corresponding Greek letter for the transform of a function denoted by a Roman letter.

2. The busy period

From the server's viewpoint the progress of a queueing system of the type considered here is a succession of alternating periods of idleness and full occupation. It is the latter which are called Busy Periods (often abbreviated to B.P's.), and the object of the investigation is to make probability statements about the lengths of B.P's.

The length in time can conveniently be treated in the manner of Bailey (1954). Essentially this consists in a study of the special process the starting point of which is the arrival and immediate reception of service of a customer, and which terminates when the server becomes idle once more. The crux of the present treatment is the simultaneous investigation of length, measured in time and number served, by an integro-difference equation technique of the type used in Conolly (1958). This is a versatile method, and the author has in essence used it also to study the B.P. in connexion with the system $GI/E_k/1$ (an extension of the present work), and with $E_k/G/1$. Earlier work by other writers is mentioned in Kendall (1951), and, more recently, there is a comprehensive study in connexion with the processes $M/G/1$ and $D/G/1$ by Takács (1955).

To be specific, in this paper we shall investigate the probability density functions $p_m(t)$ associated with the termination of a B.P. in the small time interval $(t, t+dt)$ after it began, a total of m customers having received service. Then

$$p_m = \int_0^\infty p_m(t)\,dt \tag{2·1}$$

is the probability that a total of m customers receive service during a B.P., whatever its length in time, and

$$p(t) = \sum_{m \geqslant 1} p_m(t) \tag{2·2}$$

is the P.D.F. associated with termination of the B.P. in $(t, t+dt)$, however many customers are served. In conditions favouring a steady state we shall expect to find that

$$\int_0^\infty p(t)\,dt = 1. \tag{2·3}$$

Otherwise the integral will be less than unity and a definite probability will exist of B.P's. of infinite duration.

3. Basic theory

We introduce a P.D.F. $f_{mn}(t)$ associated with the following composite event:

(i) a customer arrives at $t = 0$ and immediately begins to receive service;
(ii) a further arrival takes place in the subsequent small interval $(t, t+dt)$;
(iii) as a result of this arrival the total number waiting in the queue is $n\,(n \geqslant 1)$;
(iv) at the time t of this arrival the mth customer to be served in succession is at the counter $(m \geqslant 1)$.

Then, for $n = 1$, we have

$$f_{m1}(t) = \int_0^t \sum_{r=1}^{m-1} f_{m-r\,r}(s)\, c_r(t-s)\,ds, \tag{3·1}$$

for $m \geqslant 2$, together with

$$f_{11}(t) = c_0(t). \tag{3·2}$$

The explanation of (3·1) is that at the time of the previous arrival, in $(s, s+ds)$, say, any customer from the first up to the $(m-1)$th could have been at the counter. If it was the $(m-r)$th the previous arrival must have created a queue of r, of which the last member must be at the counter at t. The arrival at t is then the only one waiting.

Jumping directly to Laplace transforms we obtain

$$\phi_{11}(z) = \gamma_0(z), \tag{3·3}$$

and

$$\phi_{m1}(z) = \sum_{r=1}^{m-1} \phi_{m-r\,r}(z)\,\gamma_r(z), \tag{3·4}$$

where $\gamma_r(z)$ is the Laplace transform of $c_r(t)$.

Generalizing (3·4) for $n \geqslant 2$, and omitting arguments, we obtain,

$$\phi_{mn} = \sum_{r=0}^{m-1} \phi_{m-r\,n-1+r}\gamma_r. \tag{3·5}$$

Writing $$\Phi_n(y,z) = \sum_{r\geqslant 1} y^r \phi_{rn}(z)$$

for the generating function of the Laplace transforms ϕ_{mn}, we obtain by addition of (3·3) and (3·4) multiplied by appropriate powers of y the infinite difference equation

$$\Phi_1(y,z) = y\gamma_0(z) + \sum_{r\geqslant 1} y^r\gamma_r(z)\,\Phi_r(y,z); \tag{3·6}$$

and for $n \geqslant 2$, $$\Phi_n(y,z) = \sum_{r\geqslant 0} y^r\gamma_r(z)\,\Phi_{n-1+r}(y,z). \tag{3·7}$$

Following the usual method for the solution of this type of difference equation we substitute, in (3·7), x^n for Φ_n. The equation thus becomes

$$x = \Gamma(xy,z), \tag{3·8}$$

where
$$\begin{aligned}\Gamma(y,z) &= \sum_{r\geqslant 0} y^r\gamma_r(z)\\ &= \int_0^\infty a(t)\exp[-t(z+(1-y)/b)]\,dt\\ &= \alpha(z+(1-y)/b).\end{aligned} \tag{3·9}$$

If $\xi_r(y,z)$ are the roots in x of (3·8) ($r \geqslant 1$),

$$\Phi_n(y,z) = \sum_r K_r(y,z)\,\xi_r^n(y,z), \tag{3·10}$$

the numbers K being independent of n.

Now it is obvious that for $|y| \leqslant 1$

$$|\Phi_n(y,z)| \leqslant |\Phi_n(1,z)|. \tag{3·11}$$

Also $\sum_n \Phi_n(1,z)$ is the Laplace transform of the probability of an arrival while the B.P. is in progress and is certainly finite while $R(z) > 0$. Hence for $R(z) > 0$ and $|y| \leqslant 1$ the summation in (3·10) can only include ξ_r with modulus less than unity. By applying Rouché's theorem for $R(z) > 0$ we find that (3·8) has only one root ξ that fits this description, and substituting $\Phi_n = K\xi^n$ for $n \geqslant 1$ into (3·6) we obtain $K = y$. Thus, for $n \geqslant 1$, $|y| \leqslant 1$ and $R(z) > 0$,

$$\Phi_n(y,z) = y\xi^n(y,z). \tag{3·12}$$

4. The length of a busy period

Our final objective is to obtain formulae connected with the P.D.F.'s, $p_m(t)$ (cf. § 2). We shall denote by $\pi_m(z)$ the Laplace transforms of these functions, and our objective will be realized by a formula for the generating function $\Pi(y,z)$ of the $\pi_m(z)$, i.e. for

$$\Pi(y,z) = \sum_{m\geqslant 1} y^m\pi_m(z). \tag{4·1}$$

Then $\Pi(y,0)$ is the generating function of the probabilities $p_m(2·1)$ that a B.P. concludes after the service of the mth customer, while $\Pi(1,z)$ is the L.T. of the P.D.F. $p(t)$ (2·2) associated with the termination of a B.P. in $(t, t+dt)$.

We write $$A(s) = \int_s^\infty a(t)\,dt, \tag{4·2}$$

and $$d_n(t) = A(t)\,b_n(t). \tag{4·3}$$

$\delta_n(z)$ will denote the L.T. of the d_n, and its generating function

$$\Delta(y,z) = \sum_{n\geqslant 1} y^n\delta_n(z) \tag{4·4}$$

can readily be seen to be given by

$$\begin{aligned}\Delta(y,z) &= \int_0^\infty e^{-(z+(1-y)/b)s}\,A(s)\,ds\\ &= b[(1-\Gamma(y,z))/(bz+1-y)],\end{aligned} \tag{4·5}$$

since $$\int_0^\infty e^{-zs} A(s)\,ds = (1-\alpha(z))/z.$$

Now $$p_1(t) = A(t)\,e^{-t/b}/b, \tag{4·6}$$

since the probability that the B.P. terminates in $(t, t+dt)$ after the completion of service of the first customer is simply the joint probability that no one arrives in $(0, t)$ and that the service of the first customer is completed in $(t, t+dt)$. For $m \geqslant 2$ we have

$$p_m(t) = \frac{1}{b}\int_0^t \sum_{r=1}^{m-1} f_{m-r}(s)\,d_r(t-s)\,ds, \tag{4·7}$$

since, if, at the arrival before t, the $(m-r)$th customer was at the counter, this same arrival must have created a queue of $r(1 \leqslant r \leqslant m-1)$, and the whole of this queue plus the customer at the counter must be disposed of in the remaining time interval, the last in $(t, t+dt)$, during which time no further arrivals can take place. The associated L.T.'s are

$$\pi_1(z) = \delta_0(z)/b, \tag{4·8}$$

and, for $m \geqslant 2$, $$\pi_m(z) = \frac{1}{b}\sum_{r=1}^{m-1} \phi_{m-r}(z)\,\delta_r(z). \tag{4·9}$$

By comparing coefficients of powers of y it will be seen that

$$\begin{aligned}\Pi(y,z) &= [y\delta_0(z) + \sum_{r\geqslant 1} y^r \delta_r(z)\,\Phi_r(y,z)]/b \\ &= y\Delta(\xi y, z)/b \\ &= y(1-\xi)/(bz+1-\xi y).\end{aligned} \tag{4·10}$$

5. Limiting behaviour of $\Pi(y, z)$

Although (4·10) has only been proved under the conditions $|y| \leqslant 1$ and $R1z > 0$, its validity can be extended to include $z = 0$ if $|y| < 1$, and, in circumstances to be investigated, if $y = 1$. In point of fact $\Pi(y, 0)$ and $\Pi(1, 0)$ are of direct practical importance; the former since it is the generating function of the probabilities p_m relating to the length of a B.P. in terms of customers served; the latter, since it is the probability of termination in finite time. Moreover, the value N of

$$\left[\frac{d\Pi(y,0)}{dy}\right]_{y=1},$$

is the average number of customers served during a B.P.

The limiting behaviour of $\Pi(y, z)$ is required essentially for all values of the congestion index $\rho = b/a$. We commence with the study of $\Pi(1, z)$. By (4·10) this is

$$\Pi(1,z) = (1-\xi)/(bz+1-\xi), \tag{5·1}$$

where ξ is the smallest root of $$x = \Gamma(x,z) = \alpha(z+(1-x)/b). \tag{5·2}$$

If we can expand ξ into a Taylor series in z we obtain

$$\xi(z) = \sum_{n\geqslant 0} k_n z^n, \tag{5·3}$$

where $$k_n = \left[\frac{d^n\xi}{dz^n}\right]_{z=0} \Big/ n!. \tag{5·4}$$

Thus $$\Pi(1,z) = \frac{1-k_0-k_1 z - 0(z)}{bz+1-k_0-k_1 z - 0(z)}. \tag{5·5}$$

By the theory of Branching Processes it is known that:

(i) When $\rho < 1$, $k_0 < 1$. It is evident that as $z \to 0$, $\Pi(1, z) \to 1$, exactly as would be expected in steady state conditions.

(ii) When $\rho > 1$, $k_0 = 1$, and as $z \to 0$, $\Pi(1, z) \to k_1/(k_1 - b) = \rho^{-1}$. It follows that a non-zero probability exists of infinitely long B.P's.

(iii) When $\rho = 1$, $\xi = 1$ is a double root of (5·2) with $z = 0$. The expansion of $\xi(z)$ proceeds in non-integral powers, but its leading term is unity and again $\Pi(1, z) \to 1$ as $z \to 0$.

Formulae for N can be obtained from

$$\Pi(y, 0) = (1-\xi)\, y/(1-\xi y), \tag{5·6}$$

where ξ is the smallest root in x of $\quad x = \alpha[(1-xy)/b]$.

Assuming it to be legitimate, we expand ξ into a series of ascending powers of $(1-y)$, viz.

$$\xi = \sum_{n \geqslant 0} h_n (1-y)^n,$$

where

$$h_n = \frac{(-)^n}{n!}\left[\frac{d^n \xi}{dy^n}\right]_{y=1}.$$

Differentiating (5·6) and letting $y \to 1$ we obtain for $\rho < 1$

$$N = (1-h_0)^{-1}. \tag{5·7}$$

The limiting process is a little more complicated when $\rho > 1$. On the understanding that we refer only to B.P's. of finite duration we obtain

$$N = (h_2 - h_1)\rho/h_1^3. \tag{5·8}$$

6. The system $D/M/1$

It is well known that regularizing the input (which is the significance of D) has the effect of reducing customers' average waiting time as compared with a random input, and any appointment system is a recognition of the fact. We can now study its effect on a B.P. If a is the interval between successive arrivals, we have

$$\alpha(z) = e^{-az}; \tag{6·1}$$

then in connexion with numbers of customers served our interest centres on the smallest root ξ of

$$\xi = e^{-(1-\xi y)/\rho}. \tag{6·2}$$

This can be expanded into the series

$$\xi = \frac{\rho}{y} \sum_{n \geqslant 1} \frac{(ny/\rho e^{1/\rho})^n}{n(n!)}, \tag{6·3}$$

for instance, by assuming a Taylor expansion and calculating the coefficients by direct differentiation of (6·2). The result is quoted in Takács (1955).

If (6·3) be substituted into (5·6), then a little rearrangement enables the probabilities p_m concerning the termination of a B.P. after m customers have been served to be picked out as the coefficients of the powers of y. For instance,

$$p_1 = 1 - e^{-1/\rho}, \tag{6·4}$$

$$p_2 = e^{-1/\rho} - e^{-2/\rho}(1 + 1/\rho), \tag{6·5}$$

$$p_3 = e^{-2/\rho}(1 + 1/\rho) - e^{-3/\rho}(1 + 2/\rho + 3/2\rho^2), \tag{6·6}$$

$$p_4 = e^{-3/\rho}(1 + 2/\rho + 3/2\rho^2) - e^{-4/\rho}(1 + 3/\rho + 4/\rho^2 + 8/3\rho^3), \tag{6·7}$$

$$p_5 = e^{-4/\rho}(1 + 3/\rho + 4/\rho^2 + 8/3\rho^3) - e^{-5/\rho}(1 + 4/\rho + 15/2\rho^2 + 25/3\rho^3 + 125/24\rho^4). \tag{6·8}$$

These formulae can be used for all values of ρ.

For the evaluation of (5·8) we have

$$h_1 = -(\rho - 1)^{-1}, \tag{6·9}$$

and

$$h_2 = (3\rho - 2)/2(\rho - 1)^3. \tag{6·10}$$

In the following Table 1 we give values of the first five p_m for a selection of values of ρ in the two cases $M/M/1$ and $D/M/1$, together with

$$\sum_1^5 p_m, \quad \sum_1^\infty p_m \quad \text{and} \quad N.$$

It will be noticed that regularizing arrivals has the effect of increasing the probabilities of shorter B.P.'s. However, in the single instance considered with $\rho > 1$, the average number served during a finite B.P. is greater for the regular than for the random input.

Table 1. *Numerical quantities associated with the busy period for M/M/1 and D/M/1.*

ρ	p_1	p_2	p_3	p_4	p_5	$\sum_1^5 p_m$	$\sum_1^\infty p_m$	N
				M/M/1				
0·25	0·8000	0·1280	0·0410	0·0164	0·0073	0·9927	1·0000	1·3333
0·50	·6667	·1482	·0658	·0366	·0228	·9401	1·0000	2·0000
1·00	·5000	·1250	·0625	·0391	·0273	·7539	1·0000	∞
1·50	·4000	·0960	·0461	·0276	·0186	·5883	·6667	3·000
				D/M/1				
0·25	0·9817	0·0166	0·0015	0·0002	0·0000	1·0000	1·0000	1·0174
0·50	·8647	·0803	·0277	·0124	·0063	·9914	1·0000	1·2550
1·00	·6321	·0973	·0465	·0287	·0199	·8245	1·0000	∞
1·50	·4866	·0741	·0334	·0190	·0124	·6255	0·6667	4·5000

REFERENCES

Bailey, N. T. J. (1954). A continuous time treatment of a simple queue using generating functions. *J.R. Statist. Soc.* B, **16**, 288–91.

Conolly, B. W. (1958). A difference equation technique applied to the simple queue with arbitrary arrival interval distribution. *J.R. Statist. Soc.* B, **20**, 168–75.

Kendall, D. G. (1951). Some problems in the theory of queues. *J.R. Statist. Soc.* B, **13**, 151–85.

Kendall, D. G. (1953). Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain. *Ann. Math. Statist.* **24**, 338–54.

Takács, L. (1955). Investigation of waiting time problems by reduction to Markov processes. *Acta Math. Acad. Sci. Hung.* **6**, 101–28.

The design of factorial experiments: a survey of some schemes requiring not more than 256 treatment combinations

By R. G. Mitton and F. R. Morgan

British Leather Manufacturers' Research Association, Milton Park, Egham, Surrey, England

Introduction

Factorial experiments involving factors at two levels only are particularly simple to design and analyse and are widely used in consequence. Because 4- or 8-level factors can be regarded as formally equivalent to two or three factors at two levels, these can be incorporated also without trouble. The method for preparing designs for such experiments is straightforward (Fisher, 1926, 1935; Yates, 1937), but if they involve many factors the search for suitable designs may require some time. A number of designs for factorials with factors at 2, 4, and 8 levels are given in this paper in Tables 1–10. The schemes tabulated require up to 256 treatment combinations, and use is made of fractional replication (Finney, 1945), but not partial confounding. It is assumed that second and higher-order interactions can be used for the estimate of error, and for confounding (Fisher, 1935; Finney, 1947), but main effects are unconfounded and, where possible, first-order interactions also. In the schemes with fractional replication, no main effect or first-order interaction has as alias an interaction of lower order than the second.

The tables have been made as complete as possible (subject to the restrictions stated), but it may well be that some of the designs can be improved by modifications which have not been found during our search. A few of the schemes are essentially similar to those given by other workers; e.g. Yates, 1937; Brownlee, Kelley & Loraine, 1948; Cochran & Cox, 1950; and U.S. Department of Commerce, 1957.

Explanation of the tables

The standard notation developed by Fisher (1935) is used, capital letters denoting the main effects and interactions, and small letters the treatment levels. Where a scheme involves 8-level factors these are written first and allocated the first letters of the alphabet (*ABC* for the first 8-level factor, *DEF* for the second). Four-level factors are written next, and each is allocated two of the first available letters; the remaining letters (with *I* and *O* omitted) represent 2-level factors.

Each scheme is given a reference number consisting of five digits. The first three of these give the number of 8-level, 4-level and 2-level factors, respectively. Fractions of a replicate are indicated by the fourth digit m, and the number of blocks by the fifth digit n; the scheme is a $(\frac{1}{2})^m$ replicate in 2^n blocks. Where necessary, brackets are used in the scheme number; e.g. 10(11)63 indicates a scheme for one 8-level factor and eleven 2-level factors, and is a $\frac{1}{64}$ replicate in 8 blocks.

All the schemes in any one table have the same number of (equivalent) 2-level factors, and the table gives the scheme number, the generators of the defining contrasts and of the interactions confounded, the generators of the principal block, the generators of the multipliers which operate on the members of the principal block to give the remaining blocks, and the degrees of freedom available for estimating error. In several of the schemes one or more first-order interactions are confounded with blocks; these are tabulated.

The statistical analysis of the data from an experiment with fractional replication is carried out by first neglecting certain of the factors so that a complete replicate of those remaining is obtained. The contrasts are then obtained for this remainder in the usual way (Finney, 1946; Yates, 1937). All the schemes given are such that the last letters can be omitted for this purpose; thus, scheme 10(11)63 has 14 letters of which 6 are to be omitted, leaving *ABCDEFGH*.

We wish to thank the Director and Council of the British Leather Manufacturers' Research Association for permission to publish this paper.

REFERENCES

Brownlee, K. A., Kelly, B. K. & Loraine, P. K. (1948). Fractional replication arrangements for factorial experiments. *Biometrika*, **35**, 268.

Cochran, W. G. & Cox, G. M. (1950). *Experimental Designs*. New York: Wiley and Sons.

Finney, D. J. (1945). Fractional replication of factorial experiments. *Ann. Eugen., Lond.*, **12**, 291.

Finney, D. J. (1946). Recent developments in the design of field experiments. III. Fractional replication. *J. Agric. Sci.* **36**, 184.

Finney, D. J. (1947). Construction of confounding arrangements. *Emp. J. Exp. Agric.* **15**, 107.

Fisher, R. A. (1926). The arrangement of field experiments. *J. Minist. Agric.* **33**, 503.

Fisher, R. A. (1935). *The Design of Experiments*. Edinburgh: Oliver and Boyd.

Yates, F. (1937). The design and analysis of factorial experiments. *Tech. Commun. Bur. Soil Sci., Harpenden*, no. 35.

U.S. Department of Commerce (1957). Fractional factorial experiment designs for factors at two levels. *Nat. Bur. Standards Appl. Math. Series*, no. 48.

Tables 1–10

Table 1. *Designs of type* 2^5

Scheme number	Generators of defining contrasts	Generators of interactions confounded with blocks	Generators for principal block	Multipliers for other blocks	D.F. for error	1st-order interactions confounded
Full-replicate designs: 32 treatment combinations						
00502	—	*ACE, BDE*	*abe, ac, bd*	*a, b*	13	—
01302					10	—
02101	—	*ABCDE*	*ab, ac, ad, ae*	*a*	8	—
10201					6	—

Table 2. *Designs of type* 2^6

Scheme number	Generators of defining contrasts	Generators of interactions confounded with blocks	Generators for principal block	Multipliers for other blocks	D.F. for error	1st-order interactions confounded
Full-replicate designs: 64 treatment combinations						
00603	—	*ABCD, ABEF, ACE*	*acf, ade, bce*	*a, b, c*	35	—
01403					31	—
02203					27	*ABCD*
03003					23	*ABCD, ABEF, CDEF*
02202	—	*ACDE, ABDF*	*ad, bf, ce, abc*	*a, b*	30	—
03002					24	—
10302					26	—
11102					19	*ACDE*
Half-replicate designs: 32 treatment combinations						
00613	*ABCDEF*	*ACE, BDE, BCF*	*abcd, abef*	*ab, ac, ae*	6	*AB, CD, EF*
00612	*ABCDEF*	*ACD, BCF*	*ad, bf, abce*	*ab, ac*	8	*CE*
01412					5	*CE*
00611	*ABCDEF*	*ADE*	*ad, ae, bc, bf*	*ab*	9	—
01411					5	—

Table 3. *Designs of type* 2^7

Scheme number	Generators of defining contrasts	Generators of interactions confounded with blocks	Generators for principal block	Multipliers for other blocks	D.F. for error	1st-order interactions confounded
Full-replicate designs: 128 treatment combinations						
00704	—	*ABCD, ABEF, ACEG, ABG*	*abcd, abef, aceg*	*a, b, c, e*	84	—
01504					80	*ABG*
02304					76	*ABG, CDG, ABCD*
01503	—	*ADEF, BCDE, BDFG*	*abd, ace, afg, bcg*	*a, b, c*	87	—
02303					81	—
03103					74	—
10403	—	*ADE, BDF, CDG*	*ae, bf, cg, defg*	*a, c, d*	75	—
11203					67	*ADE*
12003					58	*ADE, BCFG*
11202	—	*ACDF, BEFG*	*ac, ad, be, bg, cfg*	*a, b*	70	—
12002					60	—
20102					47	*ACDF*
20101	—	*ABCDEFG*	*ab, ac, ad, ae af, ag*	*a*	48	—
Half-replicate designs: 64 treatment combinations						
00713	*ABCDEFG*	*ABCD, ABEF, ACEG*	*abcd, abef, aceg*	*ab, ac, ae*	28	—
01513					24	*ABG*
02313					20	*ABG, CDG, ABCD*
01512	*ABCDEFG*	*ACE, ABCF*	*ac, bf, dg, abde*	*ab, ad*	27	—
02312					21	—
10412	*ABCDEFG*	*ADE, BEF*	*ad, bf, cg, defg*	*ab, ac*	15	—

Table 4. *Designs of type* 2^8

Scheme number	Generators of defining contrasts	Generators of interactions confounded with blocks	Generators for principal block	Multipliers for other blocks	D.F. for error	1st-order interactions confounded
Full-replicate designs: 256 treatment combinations						
00804	—	*ABCD, ABEF, ABGH, ACFH*	*abcd, abef, abgh, acfh*	*a, b, c, e*	204	—
01604					198	—
02404	—	*ACH, AEG, BFH, DFG*	*ace, bch, bdf, afgh*	*a, b, c, d*	191	—
03204					183	—
04004					174	—
10504	—	*BEH, BFG, ABDE, ACDF*	*ad, beg, cfg, aceh*	*a, b, c, e*	183	—
11304					174	*ABDE*
12104					164	*BFG, ABDE*
20204					149	*ABDE, ACDF, BCEF*
21004					137	*ABDE, ACDF, BCEF, EFGH, BCGH*
11303	—	*BFH, BCDF, ABEFG*	*ae, ag, bf, cd, abch*	*a, b, c*	181	—
12103					170	—
20202	—	*ABDEG, ACDFH*	*ad, bg, cf, eg, fh, def*	*a, b*	158	—
21002					144	—
Half-replicate designs: 128 treatment combinations						
00813	*ABCDEFGH*	*ABCF, ABDG, ADEF*	*abcf, abdg, abeh, afgh*	*ab, ac, ae*	84	—
01613					78	—
02413					71	—
03213					63	—
10513					64	*ABCF*
11313					55	*ABCF, BCDE*
10512	*ABCDEFGH*	*ABFG, ACDF*	*af, bg, cd, eh, defg*	*ab, ac*	67	—
11312					57	—
Quarter-replicate designs: 64 treatment combinations						
00822	*ACDEG, BCDFH*	*ACF, AEH*	*ach, bdg, deh, abef*	*ae, ag*	24	—
01622					18	—

Table 5. *Designs of type* 2^9

Scheme number	Generators of defining contrasts	Generators of interactions confounded with blocks	Generators for principal block	Multipliers for other blocks	D.F. for error	1st-order interactions confounded
Half-replicate designs: 256 treatment combinations						
00914	*ABCDEFGHJ*	*ADH, AEJ, BDF, BEG*	*abde, achj, adfj, aegh*	*ab, ac, ad, cf*	195	—
01714					188	—
02514					180	—
03314					171	—
04114					162	*CGH*
04113	*ABCDEFGHJ*	*ADH, AEJ, BEG*	*abde, achj, adfj, aegh, cf*	*ab, ac, ad*	169	—
10614	*ABCDEFGHJ*	*ADJ, BDF, CFH, ACFG*	*abdg, adfh, bcef, bdej*	*ab, ac, ad, ae*	170	—
11414					160	*ACDE*
12214					150	*ACDE, ACFG, DEFG*
11413	*ABCDEFGHJ*	*AEFG, CDEF, ABDEJ*	*abcf, abde, abgh, acdg, bj*	*ab, ac, ad*	167	—
12213					155	—
20312	*ABCDEFGHJ*	*ACDG, BCEH*	*ad, ag, be, bh, fj, abcf*	*ab, ac*	141	—
Quarter-replicate designs: 128 treatment combinations						
00923	*ADEFGH, BCDEGJ*	*ADJ, BHJ, CGJ*	*beh, cfg, achj, adef*	*af, ah, bc*	75	—
01723					68	—
02523					60	—
03323					51	—
10623	*ABDEFH, BCFGHJ*	*AEF, CEH, FJ*	*afj, bdg, cdh, aceg*	*ad, ae, gj*	51	*FJ*
10622	*ABDEFH, BCFGHJ*	*AEF, CEH*	*afj, bdg, cdh, aceg, gj*	*ad, ae*	54	—

Table 6. *Designs of type* 2^{10}

Scheme number	Generators of defining contrasts	Generators of interactions confounded with blocks	Generators for principal block	Multipliers for other blocks	D.F. for error	1st-order interactions confounded
Quarter-replicate designs: 256 treatment combinations						
00(10)24	*ADEGHJ, BCFGHK*	*ACE, BCG, ADK, AGH*	*bcej, bfgh, dfjk, abchk*	*ad, ae, aj, bf,*	185	—
01824					177	—
02624					168	—
03424					158	—
04224					148	*AGH*
04223	*ABDEHJ, BCFGJK*	*CEFJ, CFHK, ACHJ*	*bdg, fhj, acdf, bcjk, cegh*	*ad, ae, ah*	155	—
05023					143	—
10723	*ACDGHJ, BCEFJK*	*ADF, BFH, BDK*	*cj, egj, adek, aefh, bghk*	*ad, ag, ah*	164	—
11523					152	—
12323					139	—
Eighth-replicate designs: 128 treatment combinations						
00(10)33	*AEFGJ, BDGHJ, BCFJK*	*ACE, CDK, EHK*	*bfg, acdj, aefhj, acghk*	*ae, ck, dh*	65	—
01833					57	—
02633					49	*CDK*
10732	*ABDEFH, BCFGHJ, ABCDFJK*	*ADF, BFJ*	*ad, bej, cgk, egh, afjk*	*bf, cj*	40	—

Table 7. *Designs of type* 2^{11}

Scheme number	Generators of defining contrasts	Generators of interactions confounded with blocks	Generators for principal block	Multipliers for other blocks	D.F. for error	1st-order interactions confounded
Eighth replicate designs: 256 treatment combinations						
00(11)34	*ACDEL, BCDHJ, CFGJKL*	*ADH, BCE, BDG, DEK*	*ahjl, acdeg, bcfgl, cefjk*	*ae, bh, fg, abd*	174	—
01934					165	—
02734	*ACGJK, EGHKL, ABDEFGJ*	*FK, CEF, DEG, EHJ*	*ahjl, abdgl, acegj, defjk*	*aj, bd, bf, acd*	156	*FK*
03534	*ABCFKL, ACEHJK, ADEFGJ*	*ACE, ADH, AFJ, BFH*	*cegl, acfhk, bdghk, abchjl*	*bl, ck, dg, ej*	145	*ABK*
01933	*AFHJK, BDFHL, CEGHKL*	*BEL, BFK, CEJ*	*akl, acgj, adgh, afhk, abcef*	*aj, bd, ce*	173	—
02733					163	—
03533					152	—
04333					140	—
10833	*BDGKL, CEHJL, ACDFGJ*	*ADF, BEG, CDL*	*af, bghj, dfhl, egjk, chjkl*	*bk, cj, dg*	149	—
11633					136	—
Sixteenth replicate designs: 128 treatment combinations						
00(11)43	*ACDEH, BCDFJ, ABCDGL, BCEGJK*	*ACJ, AEF, BCG*	*adehj, bcefj, cghjk, defkl*	*acj, adf, aeg*	54	—
01943					45	—

Table 8. *Designs of type* 2^{12}

Scheme number	Generators of defining contrasts	Generators of interactions confounded with blocks	Generators for principal block	Multipliers for other blocks	D.F. for error	1st-order interactions confounded
Sixteenth replicate designs: 256 treatment combinations						
00(12)44	*ACEHK, BCFGJ, DGJKL, FHJLM*	*AEL, AGJ, BHL, LM*	*acegk, adefj, abcfjlm, abcdefgh*	*ae, abc, adk, gjm*	163	*LM*
01(10)44					153	
02844					142	
03644					130	
04444					118	*LM, GHK*
00(12)43	*ACEHK, BCFGJ, DGJKL, FHJLM*	*AEL, AGJ, BHL*	*gjm, bchm, acegk, adefj, abcfgl*	*ae, abc, adk*	170	—
01(10)43					160	—
02843					149	—
03643					137	—
04443					124	—
10944	*ACDEFHK, BCDEFGJ, CDEJKM, CDFGHL*	*ADG, CDF, CGK, BJ*	*bjm, aeghk, cfhkl, cdeghm*	*cd, cel, acjl, ehjl*	129	*BJ, EG, FL, ACH*
11744					116	*BJ, EG, FL, ACH, ADE*
10943	*ACDEFHK, BCDEFGJ, CDEJKM, CDFGHL*	*ADG, CDF, CGK*	*bjm, ehjl, aeghk, cfhkl, cdeghm*	*cd, cel, acjl*	133	—
11743					119	—

Table 9. *Designs of type* 2^{13}

Scheme number	Generators of defining contrasts	Generators of interactions confounded with blocks	Generators for principal block	Multipliers for other blocks	D.F. for error	1st-order interactions confounded
Thirty-second replicate designs: 256 treatment combinations						
00(13)54	*ACEGK, ADHJN, ABCGJM, ACDEFM, ACHKLM*	*ACM, ADG, AEL, AFN*	*acefgj, bdeghl, cefhlmn, dfghjkn*	*acn, bjn, dfn, abdk*	149	—
01(11)54					138	—
02954					126	—
03753	*ACEGJ, BDEKN, BDFGM, ABCFHK, ABDHJL*	*ADF, BCDG, BCEFH*	*adgk, bdfm, ehjk, cdfjn, cghlm*	*acl, cek, cfg*	121	—
04553					108	*GHK*
10(10)53	*ADHMN, CEGKN, BCDEFN, BCGHJN, BCKLMN*	*ADE, AFH, BEF*	*bden, clmn, gjkl, defhmn, abcdfkl*	*acn, adf, ahj*	117	*GJ*

Table 10. *Designs of type* 2^{14}

Scheme number	Generators of defining contrasts	Generators of interactions confounded with blocks	Generators for principal block	Multipliers for other blocks	D.F. for error	1st-order interactions confounded
Sixty-fourth replicate designs: 256 treatment combinations						
00(14)63	*ACEHJ, BCFGM, ABDFHK, FGHJKL, JKLMNP, ADEFGHN*	*ACF, ADH, GP*	*achjm, dehln, glmnp, abdfjn, cfjkln*	*cfh, aefg, ahlp*	144	*GP*
01(12)63					132	
02(10)63					119	
03863					105	
04663					90	
10(11)63	*ADEFG, AGLMN, BDEJP, BEFHM, CDFKP, ABDEKL*	*ADP, BDK, JM*	*aefgp, cfghn, dfhkp, abcdhl, egjlmn*	*cjp, hmn, afkm*	99	*JM*

Confidence limits in the case of the geometric distribution

By KERMIT G. CLEMANS
U.S. Naval Ordnance Test Station, China Lake, California

Introduction

This paper presents central confidence belts for estimating m, the expected number of trials before failure, when X, the number of trials before failure, has a geometric distribution

$$\Pr(X = k) = p^k q \quad (k = 0, 1, 2, \ldots),$$

where p and q are positive constants with $p+q = 1$. The expected value of X

$$E(X) = p/q = m$$

hence $p = m/(1+m)$ and $q = 1/(1+m)$.

This model is of interest in estimating the mean life of ordnance devices that have passed acceptance tests and have an operating life throughout which the probability of failure is constant. Consider a sample of N such devices randomly selected from a population and each subjected to a series of identical use cycles until it fails. Let X_i be the number of trials before failure of the ith device. If the X_i have identical geometric distributions, a sufficient estimator of the mean m is

$$\hat{m} = \sum_{i=1}^{N} X_i/N.$$

It is further known that $\sum_{i=1}^{N} X_i$ has a Pascal distribution (Feller, 1950, § 11·3), that is

$$\Pr\left(\sum_{i=1}^{N} X_i = k\right) = \binom{N+k-1}{k} p^k q^N \quad (k = 0, 1, 2, \ldots),$$

where $p = m/(m+1)$ and $q = 1/(m+1)$. The distribution is bimodal with modes for $\Sigma X_i = m(N-1)-1$ and $m(N-1)$, $E(\Sigma X_i) = Np/q = Nm$, and $\mathrm{var}(\Sigma X_i) = Np/q^2 = Nm(m+1)$. Then $E(\hat{m}) = m$, $\mathrm{var}(\hat{m}) = m(m+1)/N$ and

$$\text{coefficient of variation of } \hat{m} = \sqrt{\left(\frac{m+1}{mN}\right)} \propto \frac{1}{\sqrt{N}}.$$

The cumulative Pascal distribution is

$$F(k) = \Pr\left(\sum_{i=1}^{N} X_i \leqslant k\right) = \sum_{j=0}^{k} \binom{N+j-1}{j} \left(\frac{m}{m+1}\right)^j \left(\frac{1}{m+1}\right)^N.$$

Construction and use of confidence belts

The belts of $1-\alpha$ confidence for each N were determined by assigning m and then finding corresponding values k_1 and k_2 where

$$F(k_1) \leqslant \tfrac{1}{2}\alpha < F(k_1+1)$$

and

$$F(k_2) < 1-\tfrac{1}{2}\alpha \leqslant F(k_2+1).$$

For the purpose of constructing curves a k' for $\frac{1}{2}\alpha$ and a k'' for $1-\frac{1}{2}\alpha$ were calculated by linear interpolation between k_i and k_i+1 and the points $(k'/N, m)$ and $((k''+1)/N, m)$ were plotted.

The summation was done by the IBM 704 Computer and the program will be kept available. Belts of different confidence for the sample sizes used in this article are available from the author. Charts I and II give two-sided 90 and 98 % central confidence limits, respectively, for m, the mean number of trials before failure, as a function of k/N, the observed mean number of trials before failure in samples of size N. The charts also give one-sided 95 and 99 % confidence limits.

In general, these confidence belts can be used in the same way as the binomial confidence belts of Clopper & Pearson (1934) or Crow's tables (1956), to estimate confidence limits or to plan in advance the size of sample necessary for a given degree of accuracy.

The confidence intervals estimated from Charts I and II are in agreement with those computed by Anscombe's method if his extra correction terms are used when necessary (Anscombe, 1948).

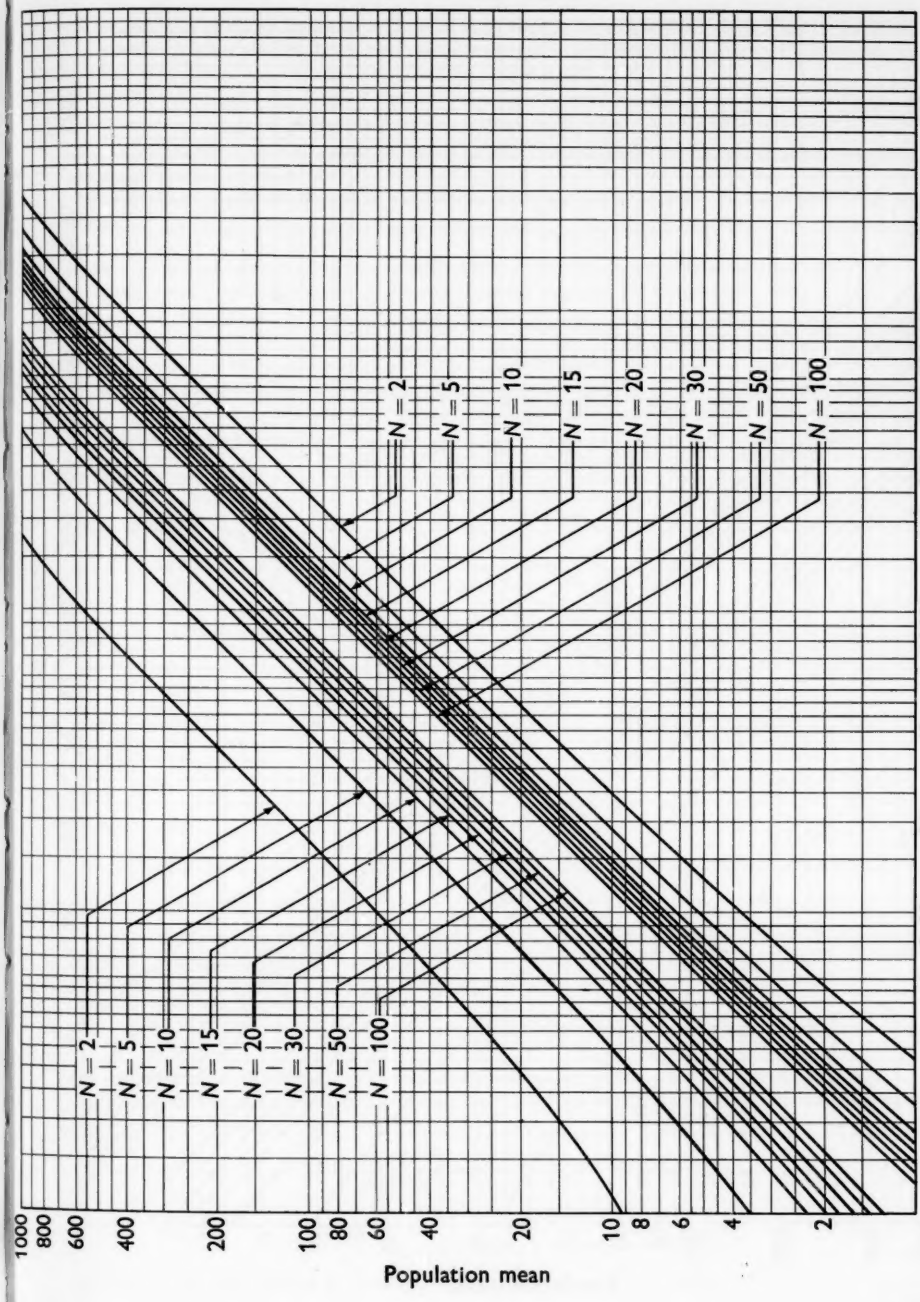
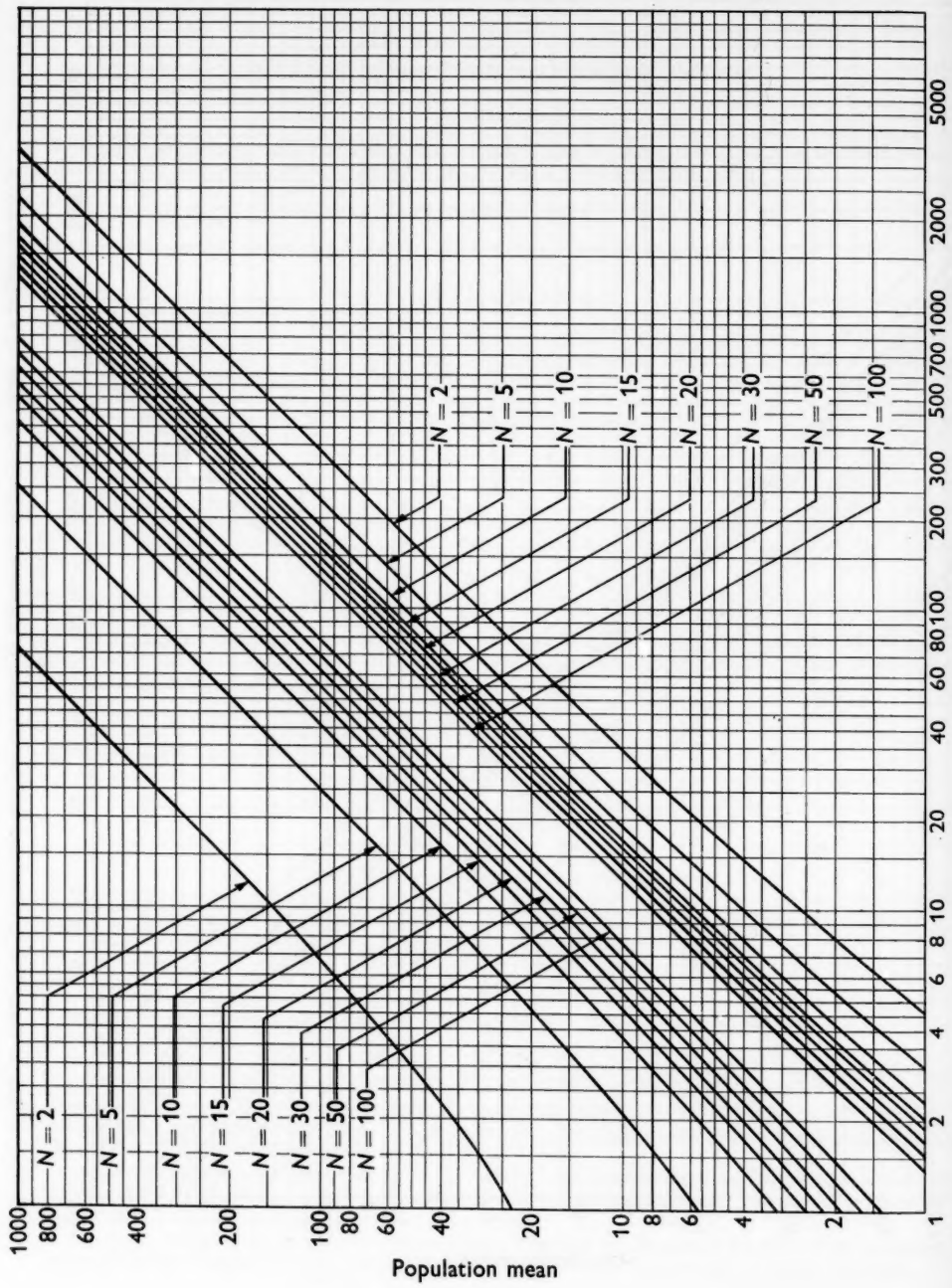


Fig. 1. Central 90 % confidence belt for the mean of a geometric distribution based on a sample of size N.


Population mean
Observed mean
1000
800
600
400
200
100
80
60
40
20
10
8
6
4
2
1
2
4
6
8
10
20
40
60
80
100
200
300
500
700
1000
2000
5000
N = 2
N = 5
N = 10
N = 15
N = 20
N = 30
N = 50
N = 100

Example 1. What are the 90 % confidence limits for m, the expected number of trials before failure, if $N = 30$ units operated for a total of $k = 15{,}000$ trials? The desired confidence limits are the ordinates of the intersection of the two curves for $N = 30$ in Chart I and the abscissa $k/N = 500$, i.e. 385–700.

Example 2. What value will the mean exceed with 95 % confidence in Example 1? The desired value is the lower limit of the central 90 % confidence band, that is 385.

Interpolation and extrapolation in the charts

Interpolation with respect to sample size may be carried out linearly on the logarithms of the population mean with arguments $N^{-\frac{1}{2}}$.

Extrapolated confidence limits for the population mean for sample size $N > 100$ are found from the equation

$$\log m = \log(k/N) + 10N^{-\frac{1}{2}} |\log m_{100} - \log(k/N)|,$$

where m_{100} is the corresponding (upper or lower) limit for sample size 100 at abscissa k/N.

Extrapolated population mean confidence belts for large sample means (k/N) are

$$m_U(N) = (1000/a)(k/N)$$

and

$$m_L(N) = (1000/b)(k/N),$$

where a and b are the abscissas at population mean 1000 for the corresponding sample size N; see Tables 1 and 2. To obtain values of a and b for values of N not tabled, interpolate linearly on values of $\log a$ and $\log b$, using $N^{-\frac{1}{2}}$ as argument.

Table 1. *Value from* 90 % *confidence belt*

N	2	3	4	5	10	15	20	30	50	100	∞
a	177·0	272·0	341·1	393·6	542·3	616·2	662·6	719·7	779·2	841·4	1000
b	2373	2100	1940	1832	1572	1460	1395	1320	1245	1173	1000
c	5·23	3·20	2·42	1·99	1·21	0·938	0·793	0·632	0·480	0·336	0

Table 2. *Values from* 98 % *confidence belt*

N	2	3	4	5	10	15	20	30	50	100	∞
a	73·57	144·8	205·3	255·4	412·7	498·2	553·9	624·6	700·5	782·1	1000
b	3324	2807	2516	2326	1883	1702	1598	1480	1366	1259	1000
c	13·29	6·55	4·47	3·49	1·89	1·42	1·18	0·925	0·696	0·484	0

Example 3. Find the size of sample necessary to ensure that the 90 % confidence interval for the population mean is not longer than 0·50 of the sample mean. This means the magnitude of the distance across the 90 % confidence belt measured along an abscissa should be $c = 0{\cdot}500$ of that abscissa. Find c in Table 1, where

$$c = 1000\left(\frac{1}{a} - \frac{1}{b}\right),$$

and a and b are the abscissas at $m = 1000$ of the confidence belt.

The closest value is $c = 0{\cdot}480$ for $N = 50$. A more accurate estimate of the sample size required may be found by interpolating between values of $N^{-\frac{1}{2}}$, using as arguments logarithms of c.

$$\frac{\log(0{\cdot}500) - \log(0{\cdot}480)}{\log(0{\cdot}632) - \log(0{\cdot}480)} = 0{\cdot}148390,$$

$$30^{-\frac{1}{2}} - 50^{-\frac{1}{2}} = 0{\cdot}041153,$$

$$N^{-\frac{1}{2}} = 50^{-\frac{1}{2}} + (0{\cdot}041153)(0{\cdot}148390) = 0{\cdot}147528.$$

This value of $N^{-\frac{1}{2}}$ is between $45^{-\frac{1}{2}}$ and $46^{-\frac{1}{2}}$; hence the sample size is taken as 46.

REFERENCES

ANSCOMBE, F. J. (1948). The transformation of Poisson, binomial and negative-binomial data. *Biometrika*, **35**, 246–54.

CLOPPER, C. J. & PEARSON, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, **26**, 404–13.

CROW, E. L. (1956). Confidence intervals for a proportion. *Biometrika*, **43**, 423–35.

FELLER, W. (1950). *An Introduction to Probability Theory and Its Applications*, **1**. New York: John Wiley and Sons, Inc.

Note regarding Statistical Research Monographs

The Institute of Mathematical Statistics and the University of Chicago have established a series of publications entitled *Statistical Research Monographs*.

The primary purpose of this series is to provide a medium of publication for material of interest to statisticians that is not ordinarily provided for by existing media. It will help fill the gap between journal articles and textbooks or treatises. Among the kinds of publications envisaged are:

New research results too lengthy for the usual journal article. In particular, authors will have ample scope for detailed exposition of their findings.

Research results of interest in both theoretical and applied statistics. At present authors of such material frequently find it necessary to publish part of their results in a theoretical journal and part in an applied journal.

Expository monographs in particular areas of statistics.

Discussions of statistical problems and techniques in particular areas of application.

Every attempt will be made to maintain the highest standards of scholarship.

The Editorial Board consists of David Blackwell (University of California), William G. Cochran (Harvard University), Henry E. Daniels (University of Birmingham), Leo A. Goodman (University of Chicago), Wassily Hoeffding (University of North Carolina), Jack C. Kiefer (Cornell University) and William H. Kruskal (University of Chicago).

Authors are invited to send manuscripts and correspondence concerning the series to Leo A. Goodman, Department of Statistics, University of Chicago, Chicago 37, Illinois.

REVIEWS

The Logic of Scientific Discovery. By KARL R. POPPER. London: Hutchinson and Co. (Publishers) Ltd. 1959. Pp. 480. 50*s*.

Prof. Popper's *Logik der Forschung* was first published in 1934. It was recognized from the first as a most important book and has had a profound effect on philosophers, though not very much on statisticians, except perhaps those who have been lucky enough to be able to attend Popper's courses on Scientific Method at the London School of Economics. Rather surprisingly, it has not previously appeared in an English translation. This new version is especially welcome at a time when scientific discovery is proceeding so rapidly and, in particular, when statisticians are still trying to put their inferential procedures on a proper formal basis.

Popper has been working on his problems since 1934, and although his views are not modified in any fundamental respect he now has much more to say. One imagines that he sat down to this translation with the idea of adding his maturer views in a few footnotes and appendices. But the new material nearly got out of control. Of the 480 pages of this book 156 are devoted to new appendices, twelve in number; and his work on what he calls the propensity interpretation of probability has been hived off into a new book (*Postscript: After Twenty Years*) which is still to appear.

To try to present the essentials of this book in a short review is perhaps ambitious; but I make the attempt in the hope that readers of this journal may be encouraged to tackle the original work.

Popper rejects the Baconian idea of induction in science, under which the study of a mass of factual observations led to the expression of a 'vintage' of universal law. He draws a sharp distinction between the process of conceiving a new idea and the examination of it. How the mind thinks of hypotheses is not his primary concern, and in any case is not a logical process in the strict meaning of the word 'logic'. He is, in this book, mainly interested in the deductive testing of a theory or a hypothesis. Further, he admits a system as scientific only if it is capable of being tested by experience. Thus the principal criterion whether a system is scientific or not is that it is testable or falsifiable.

It is not always possible to compare degrees of falsifiability, but when it is, Popper says that, of two statements, the one which is less falsifiable is the more probable. But this is a kind of probability, 'absolute logical probability', or *a priori* probability which will be new to the statistician and is not to be confused with numerical chance. The more a statement forbids, the more it says about the world of experience, and hence the greater its empirical content, its falsifiability and hence its testability; but the smaller its absolute logical probability because, to put it shortly, it forbids a larger class of possible events.

Popper then equates the *degree of falsifiability* to the *simplicity* of a hypothesis, and has some very interesting comments on the concept of simplicity itself. Thus he arrives at the schema

$$\text{testability} = \text{high prior improbability} = \text{paucity of parameters} = \text{simplicity}.$$

This is to be compared with Jeffreys and Wrinch, who effectively would have

$$\text{high prior probability} = \text{paucity of parameters} = \text{simplicity}.$$

At this point, as he himself recognizes, Popper seems to be heading for trouble. For if the parameters in a hypothesis may take an infinite set of values the prior probability of any particular hypothesis is zero; and, in fact, it was to avoid this difficulty that Jeffreys set up his method of approach to prior probability. But Popper has an answer to this. The assertion that a hypothesis has a non-zero probability is not itself testable and, whatever its metaphysical value, has no value or validity in science. We should therefore not be concerned with the probability of hypotheses but with their *degree of corroboration*. It is possible to go on learning about a theory, to test it more and more, and to increase its degree of corroboration, without increasing its logical probability, which remains at zero.

Popper then has to set up a theory of corroboration (Bewährung). In earlier work he and, following him, Carnap have used the words *degree of confirmation*, but this expression he now discards as misleading (and, in fact, as having misled some of his critics, including Carnap). Popper distinguishes between 'probability' in the sense of relative frequency, 'logical probability' in relation to statement, and 'corroboration'. In statistics he has accepted only the relative frequency interpretation of probability, though he has since modified this viewpoint by formulating his propensity theory. Within the framework of frequency theory he has in the first instance to dispose of two frequential problems: (*a*) the problem of the random sequence; this he does, as I understand it, by modifying the field of

possible tests for randomness by making it denumerable; (*b*) the problem that no assertion about a probability is testable because (again putting it roughly) anything can happen by chance; this he overcomes by an assumption rather akin to d'Alembert's, that events of small probability simply do not yield occurrences in Nature.

These difficulties behind him, he can proceed to a theory of corroboration. It is not so much the number of corroborating instances which determines the degree of corroboration as the severity of the various tests. The greater the possible severity, the greater the falsifiability; and hence the simpler hypothesis is more corroborable. The degree of corroboration of x by y is then not the same thing as the probability of x given y; and in fact it is possible to construct examples in which x_1 is strongly supported by y and x_2 is strongly undermined by it, but for which $P(x_1|y) < P(x_2|y)$.

Nevertheless, the degree of corroboration is defined as a function of probabilities. Popper's original method of doing so may, perhaps, require modification, but at least he showed that a consistent system obeying his axioms is possible. He thus measures the rationality of accepting the guess which generated the hypothesis, and this seems to be about as far as numerical evaluation of the acceptability of a theory can go.

The book contains a good deal of technical matter which will be mainly of interest to the logician, and some replies to criticism which will interest only those who are aware of the work of the original critics. Moreover, a final judgement of the system must await the sequel volume. Nonetheless, this book remains an outstanding contribution to the problem of the utilization of probabilistic ideas in the evaluation of hypotheses and will repay detailed study.

M. G. KENDALL

Experimental Designs in Industry. [Symposium held November 1956 at North Carolina State College.] Edited by V. CLEW. New York: John Wiley and Sons Inc. London: Chapman and Hall Ltd. 1958. Pp. xi + 268. 48*s*.

This book contains nine papers on experimental design read at a symposium held in November 1956 at North Carolina State College. The papers are grouped in two Parts. Part I contains four substantial papers (192 pages in all) containing, between them, most of the experimental designs which are known to have been found useful in industrial research. Part II contains five rather briefer papers on applications of experimental design techniques to specific industrial problems. The book concludes with an extensive twelve-page bibliography. There are also copious references given with some individual papers.

The outstanding feature of the first part of the book is the coherence and continuity of the four papers contained in it. This is an unusual quality in a symposium, possibly due to luck, but more probably reflecting some care in planning. In the first paper (Basic Experimental Design), the editor gives a clear, competent introduction to the general theory and application of all the simpler designs, including an account of the general linear hypothesis approach. This is followed by papers on: complete factorials, fractional factorials, and confounding (R. L. Anderson); simple and multiple regression analysis (R. J. Hader & A. H. E. Grandage) and experimental design for exploring response surfaces (G. E. P. Box & J. S. Hunter). These develop, in a natural way, the basic structure described in the first paper.

It would be quite possible for the unaided, mathematically inclined, student to obtain a good idea of modern experimental design from these four papers. They do not, of course, form an 'ideal' text-book—in particular, the theoretical bases underlying the numerical procedures are not always sufficiently clearly described for pedagogical purposes—but they do describe, in a clear unambiguous fashion, the procedures which are actually used.

The papers on specific industrial applications naturally form a rather less closely knit group. They are, however, of considerable intrinsic interest, enhanced by the fact that they give actual designs, data and analyses. The topics covered include examples of incomplete block design, fractional replication, and response surface design; and a short paper on 'Design in ordnance experimentation'.

This is a book which should prove of real value to serious students of statistics and others wishing to learn something of the techniques and outlook of modern experimental design. It is an unpretentious book (typescript with a number of surprising slips in spelling) and does not claim more for its subject than is justified. It is also quite inexpensive, by present-day standards.

N. L. JOHNSON

Analysis of Multiple Time-Series. [No. 1 of Griffin's Statistical Monographs and Courses.] By M. H. QUENOUILLE. London: Charles Griffin and Co. Ltd. 1957. Pp. 105. 24*s*.

This is the first in a series of monographs on specialized topics which are being edited by Prof. M. G. Kendall (the second was reviewed in the last issue of this journal).

Mr Quenouille's book is remarkable in that much of what it contains is new work developed over a relatively short period. The author is concerned with an extension to stationary multiple series of the correlation properties, which are well known in the case of single series and also makes a start with the associated inference problems, confining himself almost entirely to large sample theory. Specifically, the models which are considered are linear of the form $\mathbf{H}(D)\mathbf{X}_t = \mathbf{G}(D)\boldsymbol{\epsilon}_t$ where $\mathbf{H}(D)$ and $\mathbf{G}(D)$ are matrices whose elements are algebraic functions of the shift operator D, $\mathbf{X}_t$ is a vector of p elements representing the simultaneous values of p series at time t and $\boldsymbol{\epsilon}_t$ is a $p \times \infty$ matrix of uncorrelated variables extending backwards in time. One would have liked to see at the outset a statement concerning the uniqueness of this representation in the spirit of Wold's theorem for single series.

In the case where $\mathbf{G}(D) = \mathbf{I}$, the unit matrix, the correlation properties of the $\mathbf{X}_t$ are determined by the roots of the determinantal equation $|\mathbf{H}(D)| = 0$ and these may be real or imaginary in pairs. It is shown that much simplification is obtained if a transformation is made to a canonical form in which $\mathbf{H}(D)$ is diagonal with elements which are the factors of $|\mathbf{H}(D)| = 0$. In the case where these factors are real, the transformed model is that of a multivariate Markoff scheme viz. $\mathbf{A}_0\mathbf{X}_t^* + \mathbf{A}_1\mathbf{X}_{t-1}^* = \boldsymbol{\epsilon}_t^*$ and these are given special treatment in what follows.

Chapter 3 is concerned with the problem of Identification, namely, that the parameters of the scheme cannot be uniquely determined from the covariance structure. This is true even for the model cited earlier, but the situation is made more complex by the effect of superposed error. In the case of a single series this is equivalent to mixing the series with an independent random series of unknown variance. Clearly, one could make the situation more hopeless by adding an autocorrelated series with unknown autocorrelation structure, in which case it would follow that the two series could not be separated without further knowledge or specification.

Chapter 4 deals with inference problems and the methods which are developed are illustrated by means of artificially generated multiple series. Chapter 5 is devoted to miscellaneous topics and is intended to demonstrate how the analysis may be complicated in practice due to the omission of relevant variables, moving average residuals (i.e. $\mathbf{G}(D) \neq \mathbf{I}$), superposed error, non-linearities, trends and short series of observations. A very interesting method of trend analysis is proposed in which the differences of the correlations rather than those of the original series are used. Next follows chapters on estimation and goodness-of-fit tests and finally a last chapter in which a practical example (five American economic series) is subjected to the various techniques developed earlier.

It is unfair perhaps to criticize a book which explores so much new ground, but it is necessary to draw attention to the fact that it should not be read uncritically. Thus, the use of criteria of the form $\mathbf{C}_s\mathbf{C}_{s-1}$, where $\mathbf{C}_s$ is the covariance matrix corresponding to lag s, as a means of deciding on the order of the scheme is clearly unsatisfactory for moving average residuals. In the case of a single series these matrix quotients are the partial serial correlations and these are zero beyond the kth for a kth order autoregressive scheme but never zero if the residuals are moving averages. The chapter on estimation contains several surprising statements, for example, that the lower serial correlations form a sufficient set for the estimation of the parameters of a moving average scheme. Again, it is surprising that the effect of mean corrections on the serial correlations is given such little space, despite the fact that in small samples they may have considerable effect on significance tests and estimation. The relations between continuous and discrete schemes given on pp. 48–50 are interesting, but not very useful since one is rarely in a position to say that one is sampling from a 'continuous' scheme governed by a specified stochastic differential equation—in spectral analysis it is necessary to ensure that there are no oscillations between the sampling points, using some preliminary filtering where necessary.

A general point which may occur to other readers is that the approach to time-series which is pursued in this monograph is entirely empirical—this is no doubt related to the author's interest in particular subject-matter fields where no other approach seems possible. In contrast to this it is possible in electrical engineering applications to specify the $\mathbf{X}_t$ and $\boldsymbol{\epsilon}_t$ in such a way that the problem of identification disappears. Also, a linear model is used only when it is physically meaningful or else as an exploratory approach before more realistic non-linear models are used.

However, these criticisms in no way detract from the overall excellence of this monograph, an outstanding feature of which is the very elegant use of matrix methods. It is not difficult to predict that it will become a standard book worthy of serious study for years to come.

G. M. JENKINS

Fundamentals of Statistical Reasoning. [No. 3 of Griffin's Statistical Monographs and Courses.] By M. H. QUENOUILLE. London: Charles Griffin and Co. Ltd. 1958. Pp. 169. 24*s*.

'The practical man seeking a step-by-step explanation of how to carry out statistical calculations will not find it here. Nor will the mathematician find proofs of the formulae involved in deriving such procedures. If either should, however, be interested in the foundations of statistics, in the principles rather than the proofs or practices, this monograph is aimed at him.' (From the author's preface.)

After this introduction, the present reviewer was somewhat disappointed to find that the first two parts (Basic Concepts, Deductive Methods) have been devoted to an exposition of the *mathematical* concepts used in statistical theory, such as is covered by the first part of Anderson and Bancroft's *Statistical Theory in Research*. There is very little discussion of the relation of these concepts to statistical practice in the real world. The treatment is superficial and in parts inaccurate. One example must suffice: thus 'independence' (of events) is never defined, and the word is first used in the phrase 'independent exhaustive possible events of a single trial'.

A great improvement is apparent in the final part (Inductive Principles) which occupies over half the book. The Neyman–Pearson theory is sketched, and the work of Jeffreys and of Wald is mentioned, but the main emphasis is on an outline of a development of fiducial theory from axioms appropriate to a measure of belief. Some rules are given as to when joint fiducial distributions can be found and integrated; these appear to give a self-consistent theory, but there is still some difficulty in deciding whether the rules are satisfied in any given case. It is to be hoped that Mr Quenouille will continue to develop his line of approach.

Professional statisticians (who should be both mathematicians and practical men) will be stimulated by this final part, and plain mathematicians may be intrigued; but Mr Quenouille's practical men are likely still to require convincing that the theory outlined has any relevance to their problems.

C. L. MALLOWS

Multivariate Correlational Analysis. By P. H. DUBOIS. New York: Harper and Brothers. 1957. Pp. 202. $4.50.

Statisticians of the sophisticated sort are nowadays inclined to take a condescending attitude to correlation analysis and to attempt instead to operate as far as possible within the framework of regression and variance analysis. Indeed, it is hard to imagine what would have become of some of them if Fisher had never existed. Not so the psychologists, however, many of whom appear to be still contentedly following in the wake of Galton, Pearson and Yule.

The prime purpose of the present book, written for non-mathematicians, is to set out systematic methods of calculating Yule-type partial and multiple correlations. Its subsidiary purpose is to give a brief account of, and to relate to correlation analysis, the methods of factor analysis developed by Spearman and Thurstone. It is regrettable that scarcely any mention is made of orthodox modern statistics, and the distinction between population and sample is almost completely ignored. Less than a page is devoted to sampling distributions and hypothesis testing. Consequently, although the book will no doubt be of value to psychological research workers it is not likely to be very helpful to statisticians in other fields.

The book's chief merit is its demonstration of compact computing methods similar to those of Crout and Dwyer for the systematic calculation of correlation and regression coefficients by pivotal reduction of the correlation matrix. This is done with judicious mixing of partial correlation and multiple regression theory in a manner much simpler than straightforward application of Yule's theory. The author shows how partial variances and covariances may be identified at each stage of the reduction and shows how these may be combined to give correlation and regression coefficients of various kinds.

An important but irritating problem in this field is the choice of a 'best' subset of independent variables in a regression model, 'best' in the sense of maximizing the multiple correlation for subsets of a given size. For example, in psychological work we might have the results of twenty or more tests at our disposal from which we wish to select the five or six which best predict the dependent variable. It seems amazing that no one has produced an algorithm which under suitable conditions will automatically yield the best subset, apart from routine fitting of all possible regressions which will normally be out of the question because of the enormous number of possibilities involved. The customary expedient is to fit independent variables successively, taking first the one which is most highly correlated with the dependent variable, then that which in combination with the first gives the highest

regression sum of squares, and so on. This method is recommended in the book under review and a systematic computing procedure is given. As the author points out, however, not only is there no guarantee that the method will yield the best subset, it is easy to construct cases where it definitely does not. A defect of the presentation is that no rule is given for advising when to stop adding further independent variables. In this respect the presentation may be compared unfavourably with that of Wherry and that of Summerfield and Lubin (*Psychometrika*, 1951). Essentially, Wherry recommends stopping when the contribution to the regression sum of squares, due to including a further variable, is less than the estimated residual variance, while Summerfield and Lubin recommend stopping when the increment to the regression sum of squares is non-significant by the F-test. As an additional safeguard Summerfield and Lubin suggest checking that the difference between the regression sum of squares given by the reduced set of variables and the regression sum of squares given by the complete set of independent variables is small enough to be attributable to chance. These suggestions have obvious merits and it would have been helpful to have had them discussed.

Although the author's account of factor analysis is useful in that it ties it up with the earlier treatment of correlation and regression analysis, it seems altogether too short to give one a clear insight into this mysterious subject.

J. DURBIN

Introduction to Statistical Inference. By J. C. R. Li. Michigan: Edward Bros. Inc. 1957. Pp. 553. $7.50.

The publishers state that this is a non-mathematical exposition of the theory of statistics written for experimental scientists. The mathematics are in fact kept to a minimum—there is just enough to enable necessary formulae to be written down succinctly and no attempt is made at proofs. The necessary theorems are stated as articles of dogma and sampling experiments are used to illustrate their validity. In this fashion Prof. Li covers the distributions of the sample mean and variance, t, F, the variance ratio distribution, the one-way classification in the analysis of variance, randomized blocks, linear regression, analysis of covariance, factorial experiments, transformations and distribution free methods.

The level of exposition is remarkably consistent and close to the mathematical minimum aimed at throughout. The writing is clear and the ideas are set out with simplicity. The reviewer feels that the paucity of the data of experience may be a handicap to the experimental scientists, since they tend to lean heavily on actual data rather than fictitious. However this is possibly a matter of opinion, and within its limits the book can be recommended.

F. N. DAVID

Teach yourself Statistics. By R. Goodman. London: English Universities Press Ltd. 1957. Pp. 238. 7*s*. 6*d*.

As its title implies this book is an introduction to statistics for readers with a minimum of mathematics. (Elementary calculus and determinants are the most advanced techniques used and when such techniques *are* used reference to appropriate texts is often given.) There is a wealth of worked examples, but the casual reader will need a desk calculating machine by him if he is to run through them himself. The field covered is the usual basic minimum: presentation of data, measures of frequency distributions, elementary probabilities, binomial, Poisson and univariate and bivariate normal distributions, regression and correlation, rank correlation, least squares, partial correlation, sampling distributions of means, etc., elementary estimation, the t- and F-distributions, elementary testing, analysis of variance, testing in regression analysis, χ^2-tests of goodness-of-fit and contingency tables.

Whilst the topics and the order of presentation are much as in older elementary texts (Yule and Kendall springs to mind) the mode of presentation is not, at least in the earlier chapters. The author being faced with the fundamental problem in such a book, of relating the theory of probability to the practice of statistics, has gone back to an approach which would hardly have been unusual in an elementary text of the early nineteenth century. For instance, he defines probability (pp. 34–35):

'(1) If the situation is such that we can distinguish n "equally probable" outcomes of a set of circumstances C and if these are all the possible outcomes of C then if m of these outcomes entail the occurrence of an event E we say that the appropriate estimate of the probability of E given C, $p(E; C)$ is given by

$$p(E; C) = m/n.$$

(2) Otherwise, if in n occurrences of a set of circumstances C the event E has occurred m times the appropriate estimate of the probability of E given C is

$$p(E; C) = m/n.'$$

And again, if in N trials, the score x_i is obtained n_i times, $i = 1 \ldots n$ he states (on p. 43)

$$'E(x) = \lim_{N \to \infty} \left[\sum_{i=1}^{n} n_i x_i / N \right],'$$

(allowing for what must be a misprint).

Whilst such statements as these are to some extent mitigated by introductory discussion, it must be said that the arguments used are often not only loose but such as seriously and unnecessarily to confuse the student.

There is another class of loose argument also used which is due to the low level of mathematics assumed in the reader, namely 'suggestive' or heuristic mathematical proofs and these are relatively venial and in any case unavoidable. Indeed, they contribute to the reader's understanding if he is under no illusions as to their nature. In this book, however, they are often so indirect, intricate and tedious that they are of doubtful utility. Further, they use such principles as, for instance, the uniqueness theorem for moments, without even an indication that this requires proof, and so beg the question in a hidden fashion. This sort of treatment can only unnerve the intelligent reader who has a limited knowledge of mathematics and lead him to mistrust his own judgement.

The book lacks an index, which is a serious shortcoming when important principles and techniques are introduced in an *ad hoc* manner, as and when they are needed for a particular application (as for instance, criteria such as bias, efficiency, etc. of estimators or the principles of testing statistical hypotheses).

The concepts of sufficiency, maximum likelihood estimation or power of a test do not seem to be mentioned.

D. E. BARTON

Concise Tables for Statisticians. By K. C. S. PILLAI. Manila: The Statistical Center, University of the Philippines. 1957. Pp. 50. $1.50 plus postage.

This is a book of eighteen short tables on 33 octavo pages preceded by 14 pages of introductory matter. The tables fall into three groups. The first consists of seven basic tables: the normal density function and its integral, the percentage points (95 and 99 %) of the χ^2, t, F and r (the sample correlation for $\rho = 0$) and a table of random numbers. The next group consists of eight tables of analogous functions based on normal range (rather than sample S.D.). The last group contains two tables of percentage points of the largest root (table 16) and the sum of roots (table 17) of the matrix generalization of the variance ratio for multivariate analysis. In the language of Foster and Rees, table 16 consists of the percentage points (95 and 99 %) of the generalized β distribution $I_x(k; p, q)$ for $k = 2, 3, 4, 5$ at a very much coarser set of intervals than Foster's and Foster & Rees's tables. Table 17 is an approximation based on fitting Pearson curves.

The introduction, in so far as the first two groups of curves are concerned, consists of little more than specification of the function tabled, but the multivariate tables are more fully introduced with worked examples of application. No indication is directly given of the method of computation used (except for table 17 as noted) nor of what methods of interpolation are appropriate. References are given, however, to the publications in which each table (or part of it) appeared.

D. E. BARTON

Functions of Real and Complex Variables. By W. F. OSGOOD. New York: Chelsea Publishing Company. 1958. Pp. 407 + 262. $4.95.

This book—a reprint, in one volume, of two books published in 1936—deals with those branches of the theory of functions which are normally taught in a B.Sc. Special course in mathematics, and with some others, such as Fourier series, Fourier integrals, existence of solutions of differential equations, analytic continuation, Picard's theorem, and the more advanced theory of conformal mapping.

The quality of the book is, in my opinion, fairly indicated by that of the following excerpt, in which the terms 'cluster point' and 'point of condensation' mean 'limit point', and 'upper limit' means 'exact upper bound':

'THE WEIERSTRASS–BOLZANO THEOREM. *An infinite linear point set which lies in a finite interval, has at least one cluster point.*'

Consider the points x of the line, such that, if x' be any one of them, only a finite number of points of the given set (in particular, none whatever) lie to the left of x'. This point set is bounded from above. It has, therefore, an upper limit, G. This point, G, is a point of condensation. For otherwise it would be an isolated point of the set, or not belong to the set at all. In either case, there would be a point x to the right of G.

T. ESTERMANN

An Introduction to Combinatorial Analysis. By JOHN RIORDAN. New York: John Wiley and Sons Inc. London: Chapman and Hall Ltd. 1958. Pp. x + 242. 68*s*.

The great name in combinatorial analysis is undoubtedly that of the late Major Macmahon. In two books of unrivalled ingenuity and combinatory insight he delineated the theory as we know it today, knitting together the early work of De Moivre, the generalizations of Laplace and much of the elementary work in the field of number theory. The difficulty, and it is a real one, in reading Macmahon, is the very generality of the theory because although it is all there it is not easily apparent how any particular piece of theory may be applied in any given circumstance.

The work of R. A. Fisher in the combinatorial field, both in sampling moments and in experimental design, has made it essential that the worker who would research in these fields should have some knowledge of the combinatorial approach. Previously there was no alternative but to study Macmahon: Prof. Riordan has now put us all in his debt by writing a treatise which will make the approach to Macmahon less painful and to a certain extent will supplement his work.

The topics covered are permutations and combinations, generating functions, De Moivre's theorem sometimes attributed to Boole and often to Whitworth, cycles of permutations, occupancy, partitions, compositions, trees and networks, and restricted permutations. There is no tie-up with statistical methods or with probability theory. Prof. Riordan's book should, however, be made required reading for all who attempt to work in the field of so-called non-parametric methods. The standard of exposition is high and it is very easy to learn from. In the opinion of the reviewer this is an excellent treatise which fills a gap in the mathematical armour of the statisticians.

F. N. DAVID

Calculus of Finite Differences. By GEORGE BOOLE. [Edited by J. F. MOULTON.] New York: Chelsea Publishing Co. 1958. Pp. viii + 336. $4.95.

George Boole, in 1860, when he was professor of mathematics at Queen's College, Cork, wrote an elementary treatise on finite differences. After his death his book was revised by J. Moulton of Christ's College, Cambridge, in 1872. There is no record of the date of the third edition but the Chelsea Publishing Company note that the present (fourth) edition differs from the third in notational changes only.

This book, refreshingly free from the modern trend towards unnecessary mathematical rigour, has stood the test of time well. Generations of English mathematicians, statisticians and actuaries, have been reared on it, and that it has now crossed the Atlantic can only be a source of satisfaction. For those who are not familiar with 'Boole' the topics covered in the first part are Interpolation, Quadrature, Summation of Series, Bernoulli numbers, Convergency, Boole's limit for the remainder of a series. The second part is devoted mainly to difference equations. There are many examples in each chapter and a set of answers is given at the end of the book. To re-read this book after a number of years is to become aware of how little advance has been made in manipulative mathematics.

'Boole' can be recommended to students of all types and ages. The mathematical approach is simple and the amount of prerequisite learning is negligible. At the price, it is—by modern standards—in the bargain field.

F. N. DAVID

A Treatise on the Analytic Geometry of Three Dimensions. (Seventh edition.) Vol. I. By GEORGE SALMON [Edited by C. H. ROWE.] New York: Chelsea Publishing Co. 1958. Pp. xxiv + 470. $4.95.

The introduction of the geometrical method in statistics by R. A. Fisher has been followed by many derivations and proofs in distribution theory based on this type of approach. The generalization from three to n dimensions is not always skilfully or accurately done, but the fact that it is done at all makes it a necessity for the modern statistics student to have at least a nodding acquaintance with three-dimensional geometry. There are several books in current use and in the reviewer's opinion there is no reason why this, the seventh edition of Salmon's book, should be preferred to them. This being said, it must also be remarked that this is an excellent book.

The date of the original edition of this work is not stated, but it was revised in 1911 for the fifth edition and again in 1927 for the seventh edition here reprinted. Salmon, who appears to have been an inveterate writer of text-books, almost rivalling Todhunter, wrote his treatise on conic sections in the 1870's so presumably this present work will date from about that time. The topics covered

are co-ordinates and their transformation, geometrical interpretation of equations, quadriplanar co-ordinates, properties of quartics and their classification, reciprocation and duality, foci and confocal surfaces, cones, and the general theory of surfaces. One or two examples and exercises are inset at intervals in the text. The index, contrary to most text-books, is both usable and informative.

F. N. DAVID

Probleme der Statistischen Methodenlehre in den Sozialwissenschaften. (Third edition). By O. ANDERSON. Würzburg, Germany: Physica-Verlag. 1957. Pp. 358. 24.50 DM.

This is the third edition of a book first published in 1954 in the excellent series of monographs issued by the German Statistical Society. As far as can be judged, it has become a standard 'intermediate' text in the German literature.

The scope of the book follows traditional lines. There are chapters on the basic ideas of statistical methods, on measures for describing frequency distributions, on index numbers, statistical errors, probability, time series, association, correlation and so forth. As the name of the book implies, its purpose is to highlight problems, rather than to give a comprehensive account of methods. This is all to the good and makes the book more interesting to read than the standard text-book. The level of approach is largely non-mathematical and should present few problems to anyone with ordinary school mathematics. Yet the book offers more than this implies, in that the basic derivations and mathematical background for the text are given in a 100-page mathematical appendix. Thus the interest of the book is extended to a wider and mathematically more sophisticated audience.

Certain aspects of order and balance are open to criticism. The placing of the chapters on index numbers and on statistical errors so early in the book is difficult to justify. More important, it seems to the reviewer that there should have been much greater and clearer emphasis on the basic concepts of sampling theory and sample design, and on problems of statistical inference as opposed to description. Conversely, the discussion of correlation (though not of regression) might have been shorter.

But these are matters of opinion which should not obscure the real value of Prof. Anderson's text. Students—for whom the book was originally intended—and social scientists generally can hardly fail to be stimulated by its lively exposition and by the many illustrations and analogies presented in the text. A good example of Prof. Anderson's approach at its best is the discussion of cost-of-living indices: in the course of some twenty pages virtually all the main problems of index number construction are discussed with reference to what is probably the most widely publicized and misunderstood statistic there is.

One other feature of the book deserves mention. Prof. Anderson has taken the opportunity of this new edition to bring his usage of technical terms into line with that recommended for the German section of Kendall and Buckland's *Dictionary of Statistical Terms*. This is an admirable step which other authors would do well to follow. Perhaps it is not too much to hope that a growing uniformity in the use of technical terms will be followed by some standardization of statistical notation.

C. A. MOSER

Demográfia, Vol. I, Part 1. Budapest: Központi Statisztikai Hivatal. 1958. Pp. 159. 12Ft.

This new journal is mainly in Hungarian with brief summaries of articles in Russian and English. The wide gulf between Hungarian and the languages with which one is more familiar implies that by no means all statistical tables are understood.

The first paper—A Few Problems of Hungary's Demographic Policy—includes an attempt to identify the influence of recent legislative measures—the 1953 ban on the use of birth control practice, its rescinding in 1956, and the legalizing of abortion. The author stresses the urgent need to reduce the apparently large number of abortions by encouraging the use of contraception—as yet not widely employed. It appears that the original paper includes details, omitted from the English Summary, which would prove interesting to many Western demographers.

Two other papers, on the measurement of infant mortality and population projections, rest heavily on recent and now familiar U.N. work in these fields. A fourth paper deals with the questions that should be asked and tabulations made in the 1960 Hungarian census.

Five shorter articles cover: Hungary's Vital Statistics in Recent Years (crude rates and trend in number and age structure of population); Industrial Commuting in Hungary; A study of 2000 cases examined by the National Institute for Sterility Research (abortion is found a major cause of secondary sterility); Preparation for the 1959 U.S.S.R. census, and Deaths from Heart Disease.

The impression is gained that in problems of methodology, the effort of overcoming the language difficulty will not be repaid, but the situation is quite different in relation to more sociological questions involving the Hungarian way of life.

N. H. CARRIER

American Housing and its Uses. By L. WINNICK. New York: John Wiley and Sons Inc. London: Chapman and Hall Ltd. 1957. Pp. 143. 44*s*.

This statistical study of density of occupation of rooms in the U.S.A. is one of the most interesting investigations in the housing field of recent years, and Mr Winnick's scholarly treatment inspires confidence in his conclusions. The main data used is that of the U.S. Census statistics of the numbers of persons per room for households of different sizes and houses of different types and sizes. Mr Winnick measures density of occupation, *space utilization* as he calls it, by the ratio of persons to rooms, the PPR ratio; while this is, he points out, an imperfect measure for many purposes, it is the best the data permits. Information is also available about house rents and values, types of tenure and household incomes, in sufficient detail for effective examination of the relation between variations in them and in the PPR ratios. Perhaps Mr Winnick's most important conclusion from this study is that size of household tends to be more important than either rent or income in determining variations in the PPR ratio, and he finds that in consequence the range of incomes is much greater than that of the PPR ratios. A multiple correlation analysis of data for a number of cities tends to confirm his conclusion. It is of particular interest that increases in the number of rooms per person and decreases in the PPR ratios have been on the average small compared to changes in real income since the beginning of this century, though, as Mr Winnick points out, this is entirely consistent with the other results. He suggests that the great relative increase in the cost of building as well as the emergence of new durable consumer goods such as motor cars, provides some explanation of the relatively slight decline in the number of persons per room since the beginning of the century.

It is impossible to give any indication here of the many other important conclusions Mr Winnick reaches. Their interest to housing experts in this country is enhanced by the fact that the general pattern of space utilization in the U.S. in terms of PPR closely resembles that in this country. The excellent use in the inquiry of data on tenures, rents and incomes to throw light on the pattern of demand for room space demonstrates the desirability of obtaining more information about them in this country through the Census of Population.

MARIAN BOWLEY

Genetics and the Improvement of Tropical Crops. An Inaugural Lecture by Sir JOSEPH HUTCHINSON. Cambridge University Press. 1958. Pp. 30. 3*s*. 6*d*.

All who attended Prof. Hutchinson's Inaugural Lecture felt it a memorable occasion, being an exposition, by a master, of the lessons which the recent era of plant breeding can bequeath to the next phase of the art. All improvement programmes hinge on the variability of the material initially available, and Hutchinson gives a penetrating discussion of the origin of variability in wild and crop plants. Variability is inversely related to selection pressure which depends on population size and ecology. Special importance attaches to the breeding system. Variability is lost in the spread of an inbred crop plant, but is maintained if outcrossing occurs with appreciable frequency. Systems of improvement depending on a limited number of inbred lines necessarily involve early loss of genetic potentiality. The way forward, particularly in providing local races for the underdeveloped countries, would appear to lie in the interbreeding bulk or in recurrent selection. In all plant breeding work, however, as Prof. Hutchinson indicates with telling examples, there are two constant requisites for success. The selection programme must keep to a clearly formulated aim. This necessitates methods of precise statistical control. The reader may expect that plant breeding will continue to be a progenitor of new biometrical techniques.

A. R. G. OWEN

Morphological Integration. By E. C. OLSON and R. L. MILLER. The University of Chicago Press, for whom Cambridge University Press act as agents. 1958. Pp. 317. 75*s*.

This book is a collection of material by two biologists, one a palaeontologist and one a marine geologist. Profs. Olson and Miller have suggested theoretical methods for the classification and correlation of biological data and, contrary to general practice, have tried out their theory on actual observations.

The reviewer is unable to judge of the worth of their theory from a biological point of view, but from a general scientific standpoint it is sensible and constructive. Broadly speaking the authors recognize that all physical measurements on any one biological unit are correlated. (It was the existence of such correlation which led Karl Pearson to develop the theory of multiple and partial correlation some fifty years ago.) Profs. Olson and Miller discuss regression and partial correlation models, and also models where the data are considered relative to their biological function. They thus get two types of measure, the ρ-type in which correlations, etc. are found and the F-type in which are measures associated according to their biological function.

Statisticians should read this book if only to get some idea of the 'old-fashioned' view of biometry which is just now becoming fashionable again. The wealth of data is also useful from many points of view. It is also refreshing to read a book which is not a text-book of statistical theory in disguise, but one in which the elementary statistical ideas play an important but secondary role in producing a new biological theory.

F. N. DAVID

Scientific Programming in Business and Industry. By ANDREW VAZSONYI. New York: John Wiley and Sons Inc. London: Chapman and Hall Ltd. 1958. Pp. xix + 474. 108*s*.

New books on mathematical techniques of operational research follow one another so closely, even from the same publishers, that it is natural to look for distinguishing features, concerning either topics, or treatment, or both. As to topics, the book under review contains in three parts (The Fundamentals, Mathematical Programming, Programming in Production and Inventory Control) 2 chapters introduction and illustration, 4 on Linear, 1 on Convex and 1 on Dynamic Programming, 1 on the Theory of Games, and 4 on Inventory Control and Scheduling. In University lectures, the author has covered 'a good part of the text in a six-month period of 45 h of lecturing' (p. ix).

As to presentation, operational research techniques can be treated as branches of applied mathematics, with theorems, proofs, corollaries, etc., or as parts of case studies, where they gain interest from their (often spurious) realism, and lose in generality.

The author claims that 'the method of presentation...is different from that of other mathematical books....There is not a single "proof" in this book. Instead of proving theorems a great deal of effort is spent in presenting arguments so that they make sense' (p. vi).

Much as the reviewer sympathizes with such an anti-purist aim, he does not feel that the problem has been solved here. 'There are a lot of equations in the book...because they are the best way to describe the particular business problem' (p. vi). Precisely. But having been presented with notation and functional relationships, the reader is then treated to lengthy verbal explanations, and the mathematics is wasted. Also, such an approach is in danger of slipping up: we are told on p. 284 that it was proven in Chapter V that the functionals of two dual problems (in Linear Programming) have the same numerical optima, but on p. 153, where the theorem is stated, we read: 'Proof of the theorem is omitted here.' It is doubtful whether either the business manager (for whom the book is written), or the mathematician (for whom, apparently, it is not), will be satisfied. As a matter of conjecture, it is the mathematician who is better served. He can in most cases guess the essence of the missing proof, and he can appreciate the great wealth of information in this well-written book. He will also be gratified at the extensive and amazingly rapid development of a subject, aimed at practical applications, but which must be admitted—grudgingly, or with approval—to be a branch of mathematics.

Sources are mentioned in the text, but there is no comprehensive bibliography. There is, however, an author and a subject-index. The make-up of the book is excellent, and only a few misprints have been found.

S. VAJDA

Introduction to Operations Research By C. WEST CHURCHMAN, RUSSELL L. ACKOFF and E. LEONARD ARNOFF. New York: John Wiley and Sons Inc.; London: Chapman and Hall Ltd. 1957. Pp. x + 645. 96*s*.

Introduction to Operations Research is one of the most ambitious books yet written in this field. It has been developed from the contents of an annual course in operational research which is presented at the Case Institute of Technology and it is intended to be of use both to prospective customers and practitioners. The former are recommended to read certain prescribed parts of the book, whilst the latter are advised to become familiar with the whole of it. Most of the illustrative examples are taken from

industrial problems since these are, on the whole, the most accessible, and usually command the greatest interest in the audiences for whom the book is primarily designed. A warning is given, however, that the reader should not therefore assume that operational research is peculiar to the industrial scene only.

The book is divided into ten parts of which the first three and the last deal with the general structure of operational research. The remaining parts, the filling of the sandwich, are concerned with a variety of special techniques which have proved themselves to be of great use to operational research workers.

The chapters on the structure of operational research are both the weakest and the most valuable part of the book. The ideas behind the design of these chapters are admirable, but unfortunately the authors have not been able to prevent some of the exposition being wordy and diffuse. Statements such as 'It is characteristic of operations research that in the solution of each problem new problems are uncovered', and the phases of an operational research problem which are listed on page thirteen could be applied with equal validity to many disciplines and illustrate the difficulty of pinpointing in a few words the essential nature of operational research itself. Chapter 2 'An Operations Research Study of a System as a Whole', was disappointing. Although the problem outlined on p. 21 seems to require that the output rate of the main product be reduced, the end of the first stage of the research left this rate unaltered, nor is there any mention in the second stage of the research of such a reduction or, alternatively, of a step-up in sales promotion in order to keep pace with production. The dangers of suboptimization are not discussed until much later in the book. The success of the market trend analysis described on pp. 49–52 is a resounding tribute to the robust character of the methods used and the conformable behaviour of the series under examination since the level of the Type I error must frequently have been higher than the classical value of 5%. Chapters 4 and 5 on the 'Analysis of the Organization and on the Formulation of the Problem', and the excellent case study of Chapter 3, are good and succeed in emphasizing the comprehensive approach required in operational research. The final chapter on the selection of a suitable team contains a useful discussion of the desirable attributes which members of the team should possess, but it is surprising that in a subject which is so much concerned with the maximization of profits and/or the minimization of costs there appears to be no place in the team for an economist.

In parts IV–VIII, five types of problem which are encountered time and again in operational research are discussed. These are problems of inventory control, of allocation, problems concerning queues and waiting lines, problems of the replacement of perishable material and problems of competition and strategy. Most statisticians will be acquainted with these problems, since a realistic formulation of any one of them is almost sure to be in terms of random variables. Although it can be argued that the lack of mathematical sophistication in the structural chapters is an advantage, it proves a drawback in this technical section. The authors' self-imposed limitations force them into lengthy discussion and a spate of footnotes. Further, they are, for the same reason, compelled to omit many of the more interesting cases of these problems. For example, most of part IV in which inventory control is considered, is entirely deterministic; the effect of adding restrictions to, or of altering some of the parameters in, a linear programming model is carried out only in a numerical example without any general rules of behaviour or conclusions being drawn; and although the section on queuing models discusses simple queues at some length, it barely mentions the problems of queues in series, of bulk service queues, and of queues of customers with different priorities for service, etc. The chapters on testing, control and implementation of solutions which comprise part IX should be prescribed reading for all operational research workers in order to remind them that their duties extend beyond the construction of a satisfactory mathematical model.

The book is well set-up and few misprints were noticed; the worst appeared to be the wrong page number quoted in the index for the all too short discussion of the dual theorem of linear programming. The index itself is rather brief. On the other hand, an excellent list of references is quoted at the end of each chapter.

Just as most operational research workers cannot agree on a definition of operational research, so also there are many opinions on the form which an introductory text-book should take. This is an operational research problem in itself to which further study should be given. *Introduction to Operations Research* will be of use to the intelligent, sceptical executive who wishes to survey the field, and probably should be read by all operational research workers, if only for the sake of part IX. But the authors have been unable to find a level of exposition which is acceptable both to consumer and to practitioner, and one is left with the impression that the filling could have been omitted from the sandwich without seriously affecting either its food value or its flavour.

ALISON DOIG

Principles of Statistical Techniques. A First Course from the beginnings for Schools and Universities. By P. G. Moore. Cambridge: University Press. 1958. Pp. 239. 22*s*. 6*d*.

The appearance of this book is timely for those struggling with statistics in schools, with negligible direction from above in methods and application. The author has resisted the temptation to go too deeply into theory, and succeeds in making the book readable, stimulating and beneficial to anyone prepared to carry out the necessary calculations. His claim that 'this book is an attempt to put across the main principles of statistical methods to students, who are fundamentally interested in the practical applications of the subject...,' is well founded. The thirty exercises at the end of the second chapter underline his attempt to cater for all tastes. The main asset of the book is the successful use of topics selected from many fields to illustrate principles and techniques. Dr Moore has gone some way towards providing a book which ought to stimulate teachers of many subjects into the realization of the advantages of a statistical approach.

Guidance in the collection and tabulation of data, so rarely given, is nicely handled, and the pitfalls of pictorial representation well illustrated. A genuine attempt to link up probability, sampling and the binomial theorem has been made, either as a natural consequence of events or as the starting point of a theory and once again the examples are plentiful. The emphasis, perhaps by bold type, of important ideas, principles, tests or formulae would have helped, and more summaries like that at the end of Chapter 10 would have been profitable. The omission of $\Sigma(O-E)^2/E$, where O, E are observed and expected frequencies in Chapter 11, is unfortunate, but economists should find the work on indices and time series in the next chapters to their benefit. Overrating the importance of correlation is given two shocks in the last chapter; first, by the title 'Pairs of Characters' and, secondly, by the excellent example in the last paragraph of the book.

This book deserves wide use in schools, both for reference and class use, but the cost would appear to be prohibitive for the latter. A set of answers to the numerical problems would also help.

A. P. PENFOLD

OTHER BOOKS RECEIVED

The Advanced Theory of Statistics. Vol. I. Distribution Theory. By M. G. KENDALL and A. STUART. London: Charles Griffin and Co. Ltd. 1958. Pp. xii + 433. 84*s*.

Wahrscheinlichkeitsrechnung und Matematische Statistik. (Hochschulbücher für Mathematik Band 40. By M. FISZ. Berlin: Veb Deutscher Verlag der Wissenschaften. 1958. Pp. 528. 14 DM.

Planning of Experiments. By D. R. COX. New York: John Wiley and Sons. Inc. London: Chapman and Hall Ltd. 1958. Pp. 308. 60*s*.

Theory and Methods of Scaling. By W. S. TORGERSON. New York: John Wiley and Sons Inc. London: Chapman and Hall Ltd. 1958. Pp. xi + 460. 76*s*.

Linear Programming Methods and Applications. By S. I. GASS. U.S.A. and London: McGraw-Hill Book Company. 1958. Pp. 223. 50*s*. 6*d*.

Lectures on Ordinary Differential Equations. By W. HUREWICZ. New York: John Wiley and Sons Inc. and The Technology Press of the Massachusetts Institute of Technology. 1958. Pp. xiii + 122. 40*s*.

Introduction to Functional Analysis. By A. TAYLOR. New York: John Wiley and Sons Inc. London: Chapman and Hall Ltd. 1958. Pp. 423. 100*s*.

Surveys in Applied Mathematics. New York: John Wiley and Sons Inc. London: Chapman and Hall Ltd. 1958. Vol. 1. Elasticity and Plasticity. By J. N. GOODIER and P. G. HODGE. Pp. 152. 50*s*. Vol. 2. Dynamics and Non linear Mechanics. By E. LEIMANIS and N. MINORSKY. Pp. 206. 62*s*. Vol. 3. Mathematical Aspects of Subsonic and Transonic Gas Dynamics. By L. BERS. Pp. 164. 62*s*. Vol. 4. Some Aspects of Analysis and Probability. By I. KAPLANSKY, M. HALL, E. HEWITT and R. FORTET. Pp. 243. 72*s*.

The Structure of Arithmetic and Algebra. By M. H. MARIA. New York: John Wiley and Sons Inc. London: Chapman and Hall Ltd. 1958. Pp. 284. 48*s*.

Theorie des Graphes et ses Applications. By C. BERGE. Paris: Dunod. 1958. Pp. 275.

Introduccion a la Investigacion Operativa. By SIXTO RIOS *et al.* Madrid: Nuevas Graficas SA, for Instituto de Investigaciones Estadisticas del Consejo Superior de Investigaciones Cientificas. Pp. 56.

Tuberculosis in White and Negro Children. Harvard: University Press for the Commonwealth Fund. London: Oxford University Press. 1958. Vol. I. The Roentgenologic Aspects of the Harriet Lane Study. By JANET B. HARDY. Pp. 122. 60*s*. Vol. II. The Epidemiological Aspects of the Harriet Lane Study. By MIRIAM E. BRAILEY. Pp. 103. 36*s*.

Sampling Opinions. An Analysis of Survey Procedures. By F. F. STEPHAN and P. J. MCCARTHY. New York: John Wiley and Sons Inc. London: Chapman and Hall Ltd. 1958. Pp. xvi + 451. 96*s*.

Annual Epidemiological and Vital Statistics 1955. (Bilingual edition English and French.) Geneva: World Health Organization. [Available from Her Majesty's Stationery Office.] 1958. Pp. 699. £3, $12.00 or Sw.fr. 36.

Contributo allo Studio delle Tavole di Nuzialità. By G. PANIZZON. Padova, Italy: Cedam. 1958. Pp. 143.

Table of Factors for One-sided Tolerance Limits for a Normal Distribution. By DONALD B. OWEN. New Mexico: Sandia Corporation. 1958. Pp. 131.

Table of Natural Logarithms for Arguments between Five and Ten to Sixteen Decimal Places. Introduction by A. N. LOWAN. U.S. Dept. of Commerce Applied Maths. Series 53. 1958. Pp. xii + 506. $4.00.

Finite Queueing Tables. (Publications in Operations Research, No. 2.) New York: John Wiley and Sons Inc. London: Chapman and Hall Ltd. 1958. Pp. xvi + 210. 68*s*.

Table for the Solution of Cubic Equations. By H. E. SALZER *et al.* U.S.A. and London: McGraw-Hill Book Company Inc. 1958. Pp. xv + 161. 58*s*.

Table of Transport Integrals. $J_n(x) \equiv \int_0^x e^z z^n (e^z - 1)^{-2}\, dz$. (N.B.S. Circular 595.) U.S. Dept. of Commerce. 1958. Pp. 46. 40 cents.

Tables of Partitions. (Royal Society Mathematical Tables Vol. 4). By H. GUPTA, C. E. GWYTHER and J. C. P. MILLER. Cambridge University Press for the Royal Society. 1958. Pp. 39 Introduction + Pp. 132 Tables. 63*s*.

Tables for Bivariate Osculatory Interpolation over a Cartesian Grid. By H. E. SALZER and GENEVIEVE M. KIMBRO. San Diego, California: Convair-Astronautics. 1958. Pp. 40.

CORRIGENDA

(i) *Biometrika* (1950), **37**, p. 96

'On the statistical independence of quadratic forms in normal variates.' By A. C. Aitken.

line 30, *read* $AVBV$ (or $BVAV$) *for* AVB (or BVA).

(ii) *Biometrika* (1958), **45**, p. 91

'Some properties of runs in quality control procedures.' By P. G. Moore.

(*a*) In the denominator of equation (5)

read $(T+1-Tx)(1-\mu)x^{m+1}$ *for* $(T+1-\mu x)x^{m+1}$.

(*b*) In the headings of the table below equation (5)

read $1-q_m$ for q_m.

(iii) *Biometrika* (1958), **45**, p. 244

'Conditioned Markoff processes.' By W. A. O'N. Waugh.

line 6, *read* $p_i u_i \Big/ \sum_{j=0}^{\infty} p_j u_j$ *for* $p_i u_i \sum_{j=0}^{\infty} p_j u_j$.

(iv) *Biometrika* (1958), **45**, p. 327

'The properties of a stochastic model for two competing species.' By P. H. Leslie and J. C. Gower.

On the left-hand side of the first two members of equation (5·3)

read $L_1\sigma_1^2$ *for* $L^2\sigma_1^2$ and *read* $L_2\sigma_2^2$ *for* $L^2\sigma_2^2$, respectively.

TRACTS FOR COMPUTERS

Department of Statistics, University College, London

I. **Tables of the Digamma and Trigamma Functions.** By ELEANOR PAIRMAN, M.A.
Tables for summing $S = \sum_{i=1}^{\infty} \frac{1}{(p_1 i+q_1)(p_2 i+q_2)\dots(p_n i+q_n)}$ where the p's and q's are numerical factors. *Price 5s. net.*

V. **Table of Coefficients of Everett's Central-Difference Interpolation Formula.** By A. J. THOMPSON, PH.D. *Second edition. Price 7s. 6d. net.*

VIII. **Table of the Logarithms of the Complete Γ-Function (to ten decimal places) for Argument 2 to 1200 beyond Legendre's Range (Argument 1 to 2).** By EGON S. PEARSON, D.SC. *Price 5s. net.*

IX. **Log Γ (x) from $x = 1$ to 50·9 by intervals of 0·01.** By JOHN BROWNLEE, M.D., D.SC. *Price 5s. net.*

X. **On Quadrature and Cubature or on Methods of Determining Approximately Single and Double Integrals.** By J. O. IRWIN, D.SC. *Price 7s. 6d. net.*

XII. **Tables of the Probable Error of the Coefficient of Correlation.** By KARL HOLZINGER, PH.D. *Price 5s. net.*

XIII. **Bibliotheca Tabularum Mathematicarum,** being a Descriptive Catalogue of Mathematical Tables. Part I. A, Logarithms of Numbers. By JAMES HENDERSON, PH.D. *Price 9s. net.*

XV. **Random Sampling Numbers.** By L. H. C. TIPPETT, M.SC., with a Foreword by KARL PEARSON. *Price 5s. net.*

XXIII. **Tables of $\tan^{-1}x$ and $\log(1+x^2)$.** To assist in the calculation of the ordinates of a Pearson Type IV curve. By L. J. COMRIE, PH.D. *Price 5s. net.*

XXIV. **Random Sampling Numbers (2nd Series).** By M. G. KENDALL and B. BABINGTON SMITH. *Price 5s. net.*

XXV. **Random Normal Deviates.** By HERMAN WOLD. *Price 5s. net.*

XXVI. **Correlated Random Normal Deviates.** By E. C. FIELLER, T. LEWIS and E. S. PEARSON. *Price 10s. 6d. net.*

Nos. II, III, IV, VI and VII are out of print

LOGARITHMETICA BRITANNICA

A standard Table of Logarithms to Twenty Decimal Places. By A. J. THOMPSON, Ph.D. (commenced in 1922 to commemorate the tercentenary of the publication of HENRY BRIGGS'S *Arithmetica Logarithmica*).

The nine separate sections of this Table have now been issued, and the complete work consisting of the logarithms of numbers 10,000–100,000, together with Dr Thompson's General Introduction (98 pp.) is available in two bound volumes.

Price £8. 8s. 0d.

Issued by the CAMBRIDGE UNIVERSITY PRESS, Bentley House, LONDON, N.W.1
on behalf of the
DEPARTMENT OF STATISTICS, UNIVERSITY COLLEGE, LONDON
and obtainable from any bookseller

(*All rights reserved*)

BIOMETRIKA. Vol. 46, Parts 1 and 2

CONTENTS

PAGE

WALTER L. SMITH. On the cumulants of renewal processes 1
A. MERCER and C. S. SMITH. A random walk in which the steps occur randomly in time . . 30
D. J. BARTHOLOMEW. A test of homogeneity for ordered alternatives 36
C. RADHAKRISHNA RAO. Some problems involving linear hypotheses in multivariate analysis . . 49
D. N. LAWLEY. Tests of significance in canonical analysis 59
A. W. PHILLIPS. The estimation of parameters in systems of stochastic differential equations . . 67
G. E. P. BOX and H. L. LUCAS. Design of experiments in non-linear situations 77
J. D. BIGGERS. The estimation of missing and mixed-up observations in several experimental designs 91
J. H. DARWIN. Note on a three-decision test for comparing two binomial populations . . . 106
A. B. L. SRIVASTAVA. Effect of non-normality on the power of the analysis of variance test . . 114
F. N. DAVID. The z-test and symmetrically distributed random variables 123
D. H. BHATE. Approximation to the distribution of sample size for sequential tests. I. Tests of simple hypotheses 130
H. A. DAVID. Tournaments and paired comparisons 139
J. G. SAW. Estimation of the normal population parameters given a singly censored sample . . 150
K. C. SREEDHARAN PILLAI and PABLO SAMSON, JR. On Hotelling's generalization of T^2 . . . 160
ELIZABETH D. BARRACLOUGH and E. S. PAGE. Tables for Wald tests for the mean of a normal distribution 169
M. H. QUENOUILLE. Tables of random observations from standard distributions 178
E. S. PEARSON. Note on Mr Quenouille's Edgeworth Type A transformation 203
JACK LEVINE. Monomial-monomial symmetric function tables 205

MISCELLANEA

A. W. KIMBALL and E. LEACH. Approximate linearization of the incomplete β-function . . 214
D. E. BARTON and F. N. DAVID. Sequential occupancy 218
T. A. RAMASUBBAN. The generalized mean differences of the binomial and Poisson distributions 223
D. C. M. LESLIE. Determination of parameters in the Johnson system of probability distributions 229
D. E. LLOYD. Note on a problem of estimation 231
MUNUSWAMY SANKARAN. On the non-central chi-square distribution 235
K. C. SREEDHARAN PILLAI and CELIA G. BANTEGUI. On the distribution of the largest of six roots of a matrix in multivariate analysis 237
ROBERT H. SOMERS. The rank analogue of product-moment partial correlation and regression, with application to manifold, ordered contingency tables 241
B. W. CONOLLY. The busy period in relation to the queueing process GI/M/1 246
R. G. MITTON and F. R. MORGAN. The design of factorial experiments: a survey of some schemes requiring not more than 256 treatment combinations 251
KERMIT G. CLEMANS. Confidence limits in the case of the geometric distribution 260
Note regarding Statistical Research Monographs 264

REVIEWS

KARL R. POPPER. The Logic of Scientific Discovery 265
V. CLEW (Ed.) Experimental Designs in Industry 266
M. H. QUENOUILLE. Analysis of Multiple Time-Series 267
M. H. QUENOUILLE. Fundamentals of Statistical Reasoning 268
P. H. DUBOIS. Multivariate Correlational Analysis 268
J. C. R. LI. Introduction to Statistical Inference 269
R. GOODMAN. Teach yourself Statistics 269
K. C. S. PILLAI. Concise Tables for Statisticians 270
W. F. OSGOOD. Functions of Real and Complex Variables 270
JOHN RIORDAN. An Introduction to Combinatorial Analysis 271
GEORGE BOOLE. Calculus of Finite Differences 271
GEORGE SALMON. A Treatise on the Analytic Geometry of Three Dimensions 271
O. ANDERSON. Probleme der Statistischen Methodenlehre in den Sozialwissenschaften . . 272
Demográfia, Vol. I, Part 1 272
L. WINNICK. American Housing and its Uses 273
SIR JOSEPH HUTCHINSON. Genetics and the Improvement of Tropical Crops 273
E. C. OLSON and R. L. MILLER. Morphological Integration 273
ANDREW VAZSONYI. Scientific Programming in Business and Industry 274
C. WEST CHURCHMAN, RUSSELL L. ACKOFF and E. LEONARD ARNOFF. Introduction to Operations Research 274
P. G. MOORE. Principles of Statistical Techniques. A First Course from the beginnings for Schools and Universities 276

OTHER BOOKS RECEIVED 277

CORRIGENDA 279

Printed in Great Britain at the University Press, Cambridge (Brooke Crutchley, University Printer)

PAGE
1
30
36
49
59
67
77
91
106
114
123
130
139
150
160
169
178
203
205
214
218
223
229
231
235
237
241
246
251
260
264
265
266
267
268
268
269
269
270
270
271
271
271
272
272
273
273
273
274
274
276
277
279